

AN ATP-DEPENDENT PROTEOLYTIC SYSTEM
IN THE METHANOGENIC ARCHAEA

By

HEATHER LOUISE WILSON

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2001

ACKNOWLEDGEMENTS

I would like to thank Dr. Julie Maupin-Furlow for the privilege of working in her laboratory, and for her constant encouragement, helpful advice, and enthusiasm for my project. I would also like to thank the members of my committee, Drs. Henry Aldrich, Lonnie Ingram, K.T. Shanmugam, and Ben Dunn, for their thoughtful questions and advice in all aspects of my work. I am especially grateful to Drs. K.T. Shanmugam, H.C. Aldrich and L.O. Ingram and members of their labs for additional counsel and free use of material and instruments in their labs. I am greatly indebted to both Henry Aldrich and Donna Williams for their assistance and advice with electron microscopy. I am also indebted to others within the department, including Drs. James Preston, Madelaine Rasche, Tom Bobik, and members of their labs, for helpful advice and use of instruments in their labs. I thank the following persons for gifts of plasmids and/or cell cultures, as well as helpful advice: Drs. William Whitman, Warren Gardner, Brian Waters, and David Boone.

I would like to thank the members of the Maupin lab, including Mark Ou, Steve Kaczowka, LeeAnn Talarico, and Chandraraj Krishnan, for making the lab an enjoyable place in which to work. Finally, I would like to express my deep gratitude to my family for their unfailing patience, helpfulness, and willingness to listen (or at least pretend to!) when I've needed a listening ear or an encouraging word.

TABLE OF CONTENTS

	<u>Page</u>
ACKNOWLEDGEMENTS	ii
LIST OF FIGURES	v
LIST OF TABLES	vii
LIST OF ABBREVIATIONS	viii
ABSTRACT	xi
INTRODUCTION	1
LITERATURE REVIEW	3
20S Proteasomes	3
26S Proteasomes	10
AAA-type ATPases	13
ATP-dependent Proteases in <i>E. coli</i>	19
Substrate Targeting and Degradation	23
Other Proteases	29
Archaeal Proteases and Chaperones	30
Methanogens and Genetic Studies in Methanogens	40
MATERIALS AND METHODS	46
Materials	46
Strains and Media	47
DNA and Transformation Techniques	47
Protein Techniques and Enzyme Assays	49
Molecular Mass Determinations	50
Cloning the <i>M. jannaschii</i> Proteasome Genes	51
Site-directed Mutagenesis	56
Expression of the <i>M. jannaschii</i> and <i>M. maripaludis</i> <i>pan</i> Genes in <i>E. coli</i>	56
Purification of MjPAN Proteins from Recombinant <i>E. coli</i>	57
Purification of Genomic DNA from <i>M. maripaludis</i>	58
Southern Analysis of <i>M. maripaludis</i> Genomic DNA	58
Construction of a <i>M. maripaludis</i> Random Genomic λ Library	59

Isolation and Cloning of the <i>Mmp</i> Gene from <i>M. maripaludis</i>	60
Cloning of the <i>Mmp</i> Gene into <i>E. coli</i> Expression Vectors	61
Purification of MmPAN Proteins from Recombinant <i>E. coli</i>	61
Cloning of the <i>Mmp</i> and <i>Mjp</i> Genes into pWLG40 Expression Vectors	63
Synthesis of MmPAN and MjPAN Proteins in Recombinant <i>M. maripaludis</i> and Purification of C-terminal His ₆ -tagged MmPAN	63
Partial Purification of MjPAN from <i>M. jannaschii</i>	64
Purification of a 20S Proteasome and 2 NTPases from <i>M. maripaludis</i>	65
Production of Antibodies against MjPAN(Δ 1-73)	67
Non-denaturing Gel Electrophoresis of PAN Proteins	67
Electron Microscopy of Archaeal Holo-Proteasomes	67
Protein Sequence Determinations	68
DNA Sequence Determinations	68
Protein Sequence Analysis	68
RESULTS AND DISCUSSION	69
Synthesis of <i>M. jannaschii</i> 20S Proteasome and PAN in Recombinant <i>E. coli</i>	69
Analysis of <i>M. jannaschii</i> Cell-Free Extract and Partial Purification of Native MjPAN	77
Structure of Recombinant MjPAN and MjPAN(Δ 1-73)	77
Biochemical Properties of the MjPAN Complexes	80
PAN-Stimulated ATP-dependent Proteolysis of Protein Substrates by the 20S Proteasome	91
Association of MjPAN with Archaeal 20S Proteasomes	96
Proteasomal Components in <i>M. maripaludis</i>	98
Isolation of the <i>Mmp</i> Gene from a <i>M. maripaludis</i> Genomic Library	98
Purification of MmPAN from Recombinant <i>E. coli</i>	104
Partial Purification and Characterization of Two NTPases from <i>M. maripaludis</i>	106
Purification and Characterization of a 20S Proteasome (Mm20S) from <i>M. maripaludis</i>	106
Synthesis of MmPAN and MjPAN in Recombinant <i>M. maripaludis</i>	116
Purification of His ₆ -tagged MmPAN from Recombinant <i>M. maripaludis</i>	119
SUMMARY AND CONCLUSIONS	123
REFERENCES	126
BIOGRAPHICAL SKETCH	152

LIST OF FIGURES

<u>Figure</u>	<u>Page</u>
1. Predicted structure of the <i>Methanosarcina thermophila</i> 20S proteasome	5
2. Western analysis of <i>M. jannaschii</i> cell lysate.....	71
3. <i>Mjannaschii</i> PAN proteins purified from recombinant <i>E. coli</i>	72
4. Site-directed mutagenesis of the ribosomal binding site located within the <i>Mjpan</i> gene recognized in <i>E. coli</i>	74
5. Non-denaturing gel electrophoresis of MjPAN proteins	76
6. Transmission electron micrograph stereo pairs of MjPAN proteins	79
7. Effect of temperature, pH, and NaCl on the ATPase activity of MjPAN proteins.....	81
8. Nucleotide-dependent stimulation MjPAN and MjPAN(Δ 1-73) stimulation of β -casein degradation by the <i>M. thermophila</i> 20S proteasome.	92
9. ATP and AMP-PNP-mediated MjPAN stimulation of β -casein degradation by the <i>M. thermophila</i> 20S proteasome	93
10. Transmission electron micrographs of reconstituted 20S proteasome and PAN proteins.....	97
11. Deduced amino-acid sequence of the <i>Mmpan</i> gene.	100
12. Multiple sequence alignment of Archaeal PAN homologs.....	102
13. Non-denaturing gel electrophoresis of MmPAN proteins.	105
14. Q Sepharose fractions of 20S proteasome and two NTPases from <i>M.maripaludis</i> cell lysate	107
15. Non-denaturing gel electrophoresis of NTPase proteins from <i>M. maripaludis</i>	108

16. SDS-PAGE gel and Western blots of the purified 20S proteasome from <i>M. maripaludis</i>	113
17. Effect of temperature on the LLVY-Amc peptide hydrolysis activity of the 20S proteasome from <i>M. maripaludis</i>	114
18. Transmission electron microscopy of purified 20S proteasomes from <i>M. maripaludis</i>	117
19. Western blot analysis of CFE of recombinant <i>M. maripaludis</i> strains producing MmPAN	118
20. Purification of MmPAN from recombinant <i>M. maripaludis</i> using a Cu^{2+} -sepharose column	121

LIST OF TABLES

<u>Table</u>	<u>Page</u>
1. Proteases and proteolytic components found in the Archaea based on genomic sequencing.....	31
2. Chaperones and heat-shock proteins found in the Archaea based on genomic sequencing.....	36
3. Bacterial and Archaeal strains used in this study.....	48
4. Plasmids used in this study	52
5. Effect of various compounds on the ATPase activity of MjPAN and MjPAN(Δ 1-73).....	84
6. Comparison of $V_{\max}$ and K_m values for AAA-ATPases	86
7. Kinetic parameters of MjPAN and MjPAN(Δ 1-73)	87
8. Nucleotide hydrolyzing activities of MjPAN and MjPAN(Δ 1-73) proteins	89
9. The effect of dinucleotides on nucleotide triphosphate hydrolyzing activities of recombinant MjPAN and MjPAN(Δ 1-73)	90
10. Biochemical properties of NTPase #1 from <i>M. maripaludis</i>	109
11. Biochemical properties of NTPase #2 from <i>M. maripaludis</i>	110
12. Purification of a 20S proteasome from <i>M. maripaludis</i>	112
13. Multiple peptide-hydrolyzing activities of the 20S proteasome from <i>M. maripaludis</i>	115
14. Comparison of ATPase activities of cell lysate of recombinant <i>M. maripaludis</i> strains	120

LIST OF ABBREVIATIONS

A.....	Absorbance
AAA.....	ATPases Associated with a variety of cellular Activities
AMC.....	7-amido-4-methyl-coumarin
AMP.....	Adenosine 5'-monophosphate
AMP-PNP.....	5'-adenylylimido-diphosphate
ATCC.....	American Type Culture Collection
ADP.....	Adenosine 5'-diphosphate
Ap ^R	Ampicillin resistant
ATP.....	Adenosine 5'-triphosphate
β-Na.....	β-naphthylamine
BCIP.....	5-bromo-4-chloro-3-indolyl phosphate p-toluidine
BSA.....	Bovine serum albumin
CAPSO.....	3-cyclohexylamino-2-hydroxy-1-propanesulfonic acid
CL.....	Chymotrypsin-like
COG.....	Cluster of orthologous groups
CDP.....	Cytidine 5'-diphosphate
CTP.....	Cytidine 5'-triphosphate
DIG.....	Digoxigenin-11-dUTP
DMSO.....	Dimethyl sulfoxide
DTT.....	Dithiothreitol
EDTA.....	Ethylenediaminetetraacetic acid

EM.....	Electron microscopy
ER	Endoplasmic reticulum
GTP	Guanosine 5'-triphosphate
HSP	Heat shock protein
IFN	Interferon
IPTG.....	Isopropyl- β -D-thiogalactopyranoside
ITP.....	Inosine 5'-triphosphate
Km ^R	Kanamycin resistant
LB	Luria broth
MES	2-N-morpholino-ethanesulfonic acid
MHC	Major histocompatibility complex
Mj20S.....	20S proteasome from <i>M. jannaschii</i>
MjPAN.....	PAN from <i>M. jannaschii</i>
Mm20S.....	20S proteasome from <i>M. maripaludis</i>
MmPAN.....	PAN from <i>M. maripaludis</i>
MOPS.....	3-N-morpholino-propanesulfonic acid
Mt20S.....	20S proteasome from <i>M. thermophila</i>
Mt α	α subunit of the 20S proteasome from <i>M. thermophila</i>
Mt β	β subunit of the 20S proteasome from <i>M. thermophila</i>
NBT.....	Nitro blue tetrazolium
NDP.....	Nucleotide diphosphate
NEM.....	N-ethylmaleimide
NER.....	Nucleotide excision repair
Ntn.....	N-terminal nucleophile hydrolase
NSF	NEM-sensitive fusion protein

NTP	Nucleotide triphosphate
ODC	Ornithine decarboxylase
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PAN	Proteasome activating nucleotidase
PCMS	<i>para</i> -chloromercuriphenyl-sulfonic acid
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PGPH	Post-glutamyl peptide hydrolyzing
Pur ^R	Puromycin resistant
PVDF	Polyvinylidene difluoride
RP	Regulatory particle
Rpn	Regulatory particle non-ATPase
Rpt	Regulatory particle ATPase
SNAAP	Small neutral amino-acid preferring
SDS	Sodium dodecyl sulfate
SRH	Second region of homology
TES	N-tris(hydroxymethyl)methyl-2-aminoethanesulfonic acid
TL	Trypsin-like
Tricine	N-tris(hydroxymethyl)methylglycine; N-[2-hydroxy-1,1-bis(hydroxymethyl)ethyl]glycine
Tris	N-tris(hydroxymethyl) aminomethane
TTP	Thymidine 5'-triphosphate
UTP	Uridine 5'-triphosphate
X-gal	5-bromo-4-chloro-3-indolyl- β -D-galactoside

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

AN ATP-DEPENDENT PROTEOLYTIC SYSTEM
IN THE METHANOGENIC ARCHAEA

By

Heather L. Wilson

August, 2001

Chair: Julie A. Maupin-Furlow

Major Department: Microbiology and Cell Science

The 26S proteasome, composed of a 20S proteolytic core and 19S regulatory particle (RP), is responsible for energy-dependent degradation of ubiquitinated proteins in eukaryotic cells. The 20S core and proteins with similarity to the regulatory particle triphosphatase (Rpt) subunits of the RP are conserved in the Archaea.

This study reports the characterization of 20S proteasome and Rpt-like proteins from the methanoarchaea *Methanococcus jannaschii* and *Methanococcus maripaludis*. The Rpt-like protein from the hyperthermophilic *M. jannaschii* formed a 550-kDa complex which stimulated 20S proteasome-mediated protein hydrolysis in the presence of nucleotide triphosphates and was designated MjPAN (proteasome-activating nucleotidase). When compared with other ATPases, MjPAN had a high $V_{\max}$ for ATP hydrolysis [3,500 nmol P_i released $\cdot$ min $^{-1}$ $\cdot$ mg protein $^{-1}$] as well as a low affinity for ATP ($K_m=497$ μ M). Deletion of residues 1 to 73 of MjPAN did not influence complex formation or energy-dependent proteolysis. However this deletion [MjPAN(Δ 1-73)] did

increase the types of nucleotides hydrolyzed and lowered the temperature optimum for ATP hydrolysis. Based on electron microscopy, both MjPAN and MjPAN($\Delta 1-73$) associated with the ends of the 20S proteasome, resembling proteasome complexes isolated from eukaryotes. These results suggest that archaeal 20S proteasomes require a nucleotidase complex such as PAN to mediate energy-dependent proteolysis of folded proteins in the cell.

Methanococcus maripaludis is a mesophilic methanoarchaeon for which genetic tools and high-yield culture techniques have been developed. Two protein complexes with NEM-inhibitable nucleotidase activity similar to MjPAN were identified in *M. maripaludis*; however, neither protein cross-reacted with anti-MjPAN antibody. Thus, a *pan* gene, which encoded a protein with 70% identity to MjPAN, was isolated from *M. maripaludis*. The MmPAN protein, which was purified from recombinant *Escherichia coli*, cross-reacted with an anti-MjPAN antibody and formed a 600-kDa complex. The *Mjpan* and *Mmpan* genes were also synthesized in *M. maripaludis*. This increased the level of PAN protein in the methanogen as detected by Western blot and ATPase assays of cell lysate. Current and future work involves using the constructs developed in this study to further characterize PAN proteins and to identify cellular components that interact with PAN and the 20S proteasome.

INTRODUCTION

Proteolysis is now understood to be an important part of the "life-cycle" of a protein from the time it is first synthesized on the ribosome, to the time the protein is degraded and its amino acids are recycled to be re-used in biosynthetic processes by the cell. In eukaryotes, proteolysis of regulatory proteins such as cyclins has been shown to be required for timely progression through the cell cycle (reviewed in 167, 246) and is also essential in degradation of foreign proteins for antigen presentation during the immune response (59). Proteolysis of specific transcription factors and cell-division proteins is also required in eubacterial cells for cell division and sporulation (reviewed in 78, 177). In contrast with eukaryotic cells, prokaryotic cells lack membrane-encased organelles, such as lysosomes, for segregation and specific degradation of cellular proteins. Therefore, the necessity for proteases which separate their active sites away from cytosolic proteins is especially important in these organisms.

Theoretically, the primary amino acid sequences of proteins contain all of the information necessary for proper folding (8). However, in the crowded environment of the cell, which has been estimated to contain at least 30% w/v protein (152), both chaperones and proteases are needed to mediate protein folding and, when necessary, protein unfolding and degradation (53). On the surface, protein folding and assembly would appear to be completely different processes than protein unfolding and degradation. However, it is likely that some of the same proteins involved in degradation of protein substrates can also act as chaperones to aid in the correct folding of proteins

(reviewed in 239, 300). Obtaining an understanding of how substrates are specifically partitioned between either folding or proteolysis and determining the mechanisms involved in maintaining the proper balance of protein folding and degradation in the cell have become the focus of numerous studies (reviewed in 80).

Archaea reside in a wide variety of often extreme environments; therefore their mechanisms for dealing with stress, and especially their proteolytic and chaperone systems, are of particular interest (reviewed in 154, 158). In this study, the proteasome and the proteasome-activating nucleotidase (PAN) were studied in two archaea, *Methanococcus jannaschii* and *Methanococcus maripaludis*. Homologs of the proteasome and PAN are well conserved in both eukarya and archaea, but are not as well conserved in the eubacteria. Archaea are similar to both eubacteria (in their metabolic processes and ultrastructure) and eukarya (in their transcription and translation processes). Therefore, study of the proteolytic system found in these methanogens and other Archaea should add to an overall understanding of ATP-dependent proteolysis in all three phylogenetic domains.

The overall objective of this project was to obtain a better understanding of the ATP-dependent proteases, especially in the methanogenic archaea, with special emphasis on the hyperthermophile *Methanococcus jannaschii* and the mesophile *Methanococcus maripaludis*.

LITERATURE REVIEW

This review of the literature will include an overview of proteolytic and chaperone systems in eukaryal, archaeal, and eubacterial cells, with a particular emphasis on the proteasome. It will also include a survey of the proteolytic and chaperone machinery present in the archaea, based on genomic sequencing as well as direct experimentation. It will conclude with a review of genetic studies which have been performed in archaea, with a special emphasis on genetic work done in the methanogens *Methanococcus voltae* and *Methanococcus maripaludis*.

20S Proteasomes

The proteasome core, with a sedimentation coefficient of 20S, is a barrel-shaped, four-ringed structure which is highly conserved in both eucarya and archaea, but among eubacteria has been found only in the gram-positive actinomycetes (243). The 20S proteasome was first identified in erythrocyte cells in 1968 (91) and since then has been isolated from numerous organisms and given over 20 different names, including macropain (161), multicatalytic proteinase complex or MCP (176), and low-molecular-weight polypeptide complex or LMP (164). The name proteasome was first used in 1988 (10) and has since become the most widely used name for this complex. The proteolytic active sites of the proteasome are sequestered within an internal chamber, which prevents the active sites from coming into contact with cellular proteins. Substrate proteins can only enter the protease by first passing through a very small opening formed by the outer rings on either end of the protease. Therefore, only completely unfolded proteins and

peptides can enter the proteasome, where they are then degraded in an ATP-independent reaction (discussed below).

Archaeal 20S Proteasomes. The 20S proteasome from *Thermoplasma acidophilum* (41), a thermoacidophilic archaeon, was the first archaeal 20S proteasome which was biochemically characterized or produced in recombinant *E. coli* (302). Since then, 20S proteasomes have been purified from several other archaea including *Methanosarcina thermophila* (157), *Pyrococcus furiosus* (13), and *Haloferax volcanii* (286). In addition, each archaeal genome which has been sequenced contains ORFs predicted to encode the two (α and β) subunits of 20S proteasomes. Most archaeal proteasomes are presumed to contain seven α or β subunits per ring. The α subunits, which form the outer rings of the four-ring complex, limit access of the substrate into the inner catalytic chamber formed by two rings of seven β subunits each (Fig. 1).

Crystal structure of the *T. acidophilum* 20S proteasome. The crystal structure of the *T. acidophilum* 20S proteasome has been solved to a resolution of 3.4Å (151). The overall dimensions of this proteasome are 148Å by 113Å. The α and β subunits are quite similar in their tertiary structure and crystallographic fold, each consisting of two five-stranded antiparallel β sheets with flanking α helices, even though they have low primary sequence identity (<20%). The α subunits contain an additional 35 amino-acid long α helix near the N-terminus, and the N-terminal 12 residues are disordered and not visible in the crystal structure. Seven α subunits assemble into a ring with an opening of only 13Å in diameter. This opening allows only completely unfolded substrates to enter the proteasome complex and come in contact with the proteolytically active β subunits.

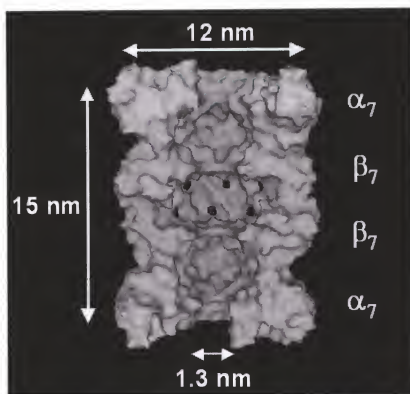


FIG. 1. Predicted structure of the *Methanosarcina thermophila* 20S proteasome. The image is an atomic model based on the crystal structure coordinates of 20S proteasomes from *Thermoplasma acidophilum* and *S. cerevisiae*. The active sites involved in peptide bond hydrolysis are indicated in black. Computational results were obtained using software programs from Molecular Simulations Inc. (<http://www.msi.com>) and from the Swiss Institute of Bioinformatics (<http://www.expasy.ch>).

The interface between the α_7 and β_7 rings creates an opening which is 27 Å in diameter, and the two β_7 rings form a central cavity which is 53 Å in diameter. Based on the crystal structure, as well as other biochemical inhibitor and mutagenesis studies, the 20S proteasome was determined to be an N-terminal threonine protease (151, 156, 212). Each β subunit has one active site as demonstrated by binding of the inhibitor N-acetyl-Leu-Leu-norleucinal (Ac-LLnL-al) to the active site N-terminal threonine residues in the crystal structure, and the distance between each active site is approximately 28 Å. The N-terminal threonine serves as a proton acceptor and nucleophile, with glutamate¹⁷ and lysine³³ forming a salt bridge to complete the active site. Aspartate¹⁶⁶ is also likely to be involved, as site-directed mutagenesis of this residue inactivates the proteasome without perturbances in the overall structure (211).

The Ntn-hydrolase family. The Ntn (N-terminal nucleophile)-hydrolase family of proteases uses the side chain of the N-terminal amino acid as a catalytic nucleophile during proteolysis (23). All Ntn-hydrolases are synthesized as a precursor in which the N-terminal residue, which can be a cysteine, serine, or threonine, is exposed by autocatalytic processing and removal of a propeptide. Even though these proteins often have no detectable primary sequence identity, all Ntn-hydrolases which have been crystallized have a similar β -sandwich fold consisting of two five-stranded antiparallel β sheets with α helices on either side (19, 49, 84, 151). This fold is essential for orientation of the active site residues including the N-terminal nucleophile as well as acidic and basic residues (Glu¹⁷ and Lys³³ in the *T. acidophilum* proteasome) which behave as a salt bridge to complete the catalytic triad (19, 49, 84, 151). Most 20S proteasomes and the eubacterial homolog HslUV (discussed below) are Ntn-hydrolases

with N-terminal threonine residues. Changing the N-terminal threonine of the *T. acidophilum* 20S proteasome to serine has little effect on the cleavage of the amide bond of small fluorogenic substrates (156, 212). However, the mutant proteasome was less active during the cleavage of full-length proteins (127) and was not as stimulated as the native 20S proteasome by PAN (proteasome-activating nucleotidase, discussed below) in the degradation of FITC-labeled protein substrates (127).

Eukaryotic 20S proteasomes. Eukaryotic 20S proteasome cores are more complex than archaeal proteasomes, with seven different subunits forming each ring for a total of 14 different subunits per 20S complex. However, each individual subunit can be classified as either an α -type or a β -type subunit, based on sequence similarity to the subunits of the archaeal proteasome. Both the archaeal and eukaryotic proteasomes contain seven α -type subunits form the outer rings, and two rings of seven β -type subunits create the inner catalytic chamber. Interestingly, eukaryotic 20S proteasomes have only three N-terminal threonine-containing active β -type subunits (β_1 , β_2 , and β_5) per ring, for a total of six active sites in the full 20S complex. In antigen presenting cells of higher eukaryotes, these β subunits can be replaced by three other active subunits (LMP2/ β_1 , LMP7/ β_2 , and MECL1/ β_5) in the presence of interferon- γ (IFN- γ) to form the immunoproteasome (1, 56, 58, 168). Notably, LMP2 and LMP7 are encoded by the major histocompatibility complex (MHC) class II locus (229), and cells deficient in these two subunits demonstrate defects in antigen presentation (59). Differences observed when these subunits are incorporated into the 20S proteasome include a change in the peptide cleavage specificity, with a decrease in cleavage after acidic residues (trypsin-like activity) and an increase in cleavage after basic and hydrophobic residues (46). It has

been determined that the majority (>73%) of peptides bound in the MHC complex have hydrophobic or aromatic residues at their C-termini (169).

Crystal structure of the 20S proteasome of yeast. The crystal structure of the 20S proteasome from *S. cerevisiae* has been solved to a resolution of 2.4Å (84). Unlike the *T. acidophilum* 20S proteasome in which the N-terminal residues of the α subunits were disordered, forming a 13Å opening, the N-terminal residues of the yeast proteasome could be resolved and appeared to completely close the ends of the cylinder. This was not surprising, since proteasomes purified from eukaryotes are usually inactive and must be treated with either heat or detergents such as SDS for hydrolysis of peptide substrates (175, 222). *In vivo*, eukaryotic 20S proteasomes are presumed to be stimulated to full activity by the binding of regulatory components [such as the 19S cap/regulatory particle (RP)] to the ends of the cylinder (97). In addition, each of the 14 different α and β subunits were found to be in a defined place in the yeast 20S proteasome crystal structure, as had been previously predicted based on studies using immunoelectron microscopy (134, 135). Notably, only the three active β subunits contained the inhibitor Ac-LLnL-al bound to the N-terminal threonine. In addition, the other residues essential for activity in the archaeal proteasome (Glu¹⁷, Lys³³, and Asp¹⁶⁶), are also conserved in these three active subunits of the yeast 20S proteasome (84).

Eubacterial proteasomes. Proteasomes have also been purified from gram-positive actinomycetes, one of which (from *Rhodococcus* sp.) is of intermediate complexity, having two of each type of subunit (243). While the 20S proteasome has not been found in gram-negative eubacteria, the double-ringed HslV complex has been termed the “prokaryotic proteasome”, based on its N-terminal threonine active site and the similarity

of the crystallographic fold of its subunits to the β subunits of the 20S proteasome (discussed below).

20S proteasome assembly. The model for both eukaryotic and archaeal proteasomes is that the α subunits of the complex form a ring which provides a scaffolding upon which the β subunits are assembled (30, 31, 96). Two half-proteasomes formed of one α ring and one β ring then assemble into a full 20S proteasome. Archaeal α subunits, when synthesized alone, form both single and double seven-membered rings (156, 287, 301). The N-terminal propeptides of the β subunits are removed autocatalytically to expose an active-site threonine during assembly of the 20S complex. If β subunits are synthesized alone, the propeptide is not removed and the subunits generally remain as inactive monomers, dimers, and trimers (156, 287, 301). While the β subunit propeptide may serve an essential role in the archaeal cytosol, the propeptide is not required for assembly of archaeal 20S proteasomes in recombinant *E. coli*. Synthesis of the β -subunit without the propeptide and subsequent mixing of equimolar concentrations of the α and β subunits results in the assembly of fully functional 20S proteasome complexes (156, 287, 301).

Some of the proteasome α -type subunits from eukaryotes can also form single and double rings when synthesized alone (66, 67). The three active (N-terminal threonine-containing) β -type subunits are also synthesized as inactive precursors, and the prosequence is necessary for assembly of the 20S proteasome in the eukaryal cell (30). In addition to mediating proper 20S proteasome assembly, the propeptides also prevent N^α-acetylation and inactivation of the N-terminal catalytic threonine in yeast (9). When the β -type subunit Pre2/Doa3 (β 5) was expressed without the prosequence in yeast, this

subunit was not incorporated into the proteasome (31). However, the prosequence could operate in *trans*, indicating a possible chaperone-like function (31). In addition, Ump1, a small protein of 16.8 kDa, has also been identified as a chaperone which interacts specifically with pre-processed β subunits in the half-proteasome and is necessary for proper assembly of the 20S proteasome in yeast (194). The Ump1 protein is internalized and rapidly degraded by the proteasome during assembly (194). Transcription of *ump1* is also increased in mammalian cells in the presence of IFN γ , indicating a possible additional role for Ump1 in the formation of the immunoproteasome (28).

The proteasome from *Rhodococcus* sp. has a different assembly pathway than either eukaryal or archaeal proteasomes (160, 298). The α subunits, when expressed alone, do not form a ring. Instead, the α and pre-processed β subunits form a dimer, which then assembles into a half-proteasome made up of a double ring of α and β subunits (160, 298). The half-proteasomes then assemble into a full 20S complex during which the propeptides of the β subunits are removed (160, 298).

26S Proteasomes

The 2-MDa 26S proteasome complex, which is composed of the 20S proteolytic core as well as an RP with a sedimentation coefficient of 19S, has been found thus far only in eukaryotic cells (reviewed in 40, 166, 246, 247). It was first identified in rabbit reticulocytes based on its ability to degrade substrates which had been labeled with ubiquitin (101, 102), and subsequently it has been shown that the majority of cytoplasmic proteins are degraded by the 26S proteasome in eukaryotes (94, 269). In addition, the 26S proteasome appears to be the primary proteolytic machinery responsible for degrading foreign proteins into peptides for presentation in the MHC Class I complex (198).

Numerous proteins have been identified as substrates of the 26S complex, and many of them are very short-lived regulatory proteins, such as cyclins, transcription factors, and cyclin-dependent kinase (cdk) inhibitors (reviewed in 167, 246). Notably, the 26S proteasome is the only two-component ATP-dependent protease identified in eukaryotic cells, in distinct contrast to eubacterial cells, which contain several (discussed below).

The regulatory particle (RP) of the eukaryotic 26S proteasome. The regulatory particle or RP, which has also been called the 19S cap (188), PA700 complex (42), or the μ particle (129), is composed of ~15 subunits ranging in size from 25 to 110 kDa. The RP associates with 20S proteasomes in an ATP-dependent manner and is responsible for the recognition and translocation of substrates into the proteolytic 20S core of the proteasome (reviewed in 40, 166, 246, 247). The RP can be further divided into Base and Lid sub-domains (74). The Base sub-domain, which directly contacts the ends of the 20S proteasome, contains eight subunits, six of which have been demonstrated to be members of the AAA (ATPases associated with a variety of cellular activities) family of proteins (42, 43). These AAA proteins of the Base sub-domain are predicted to form six-membered rings which interact directly with the seven-membered α -rings of the 20S proteasome in an asymmetrical mismatch. As discussed previously, eukaryotes contain seven different α subunits in the outer rings of the 20S proteasome, which are all in a well-defined place within the ring (134), and the six AAA components of the Base sub-domain of the RP have also been shown to interact specifically with one another (196) in a defined place within the complex (74). It is thus likely that ATPases of the RP interact with specific α subunits of the 20S core, albeit in an asymmetrical mismatch. Proposed functions of the ATPase components include coupling of ATP hydrolysis to unfolding

and feeding of substrate molecules into the 20S core of the proteasome (72, 73). In addition, there is evidence that the ATPases of the Base may open the normally closed gate formed by the α subunits in the yeast proteasome (83). The Lid sub-domain of the RP also contains several subunits that have been characterized. Several of the Lid subunits contain coiled-coil or KEKE motifs, which have been implicated in protein-protein interactions (196, 275). In general, however, less is known about the function of the subunits of the Lid, as there is little sequence conservation among them.

Nomenclature of proteins in the RP. Several different systems have been used to name the subunits in the RP, including numbering the 15 subunits originally identified in the *S. cerevisiae* RP as S1 to S15 (47). However, as more subunits were identified from several different organisms, the names became more ambiguous and confusing. In addition, it is now becoming clear that the RP is not a well-defined complex, but rather that the subunit composition of the RP may change depending on the specific needs and conditions of the cell (25). Many of the regulatory proteins were originally identified from unrelated studies of yeast mutants, and it was not known that these proteins were also members of the RP. Thus, the original names for these subunits often contained no indication that they were members of the RP of the 26S proteasome. Therefore, a "unified nomenclature" was developed for the proteins in the RP (55). The six ATPase members of the Base sub-domain of the RP were renamed Rpt1p to Rpt6p, for regulatory particle triphosphatase proteins (55). For example, Rpt2p is the designation for the subunit previously identified as S4 in humans, p56 in bovine cells, and Mts2, Yhs4p, or Yta5 in yeast; and Rpt6p is the designation for the subunit previously identified as S8 in humans, p45 in bovine cells, and Sug1, Cim5, or Cr13 in yeast. Mutants defective in each of the

six Rpt proteins can show widely variant phenotypes, which helped to contribute to the diverse nomenclature in the literature even within the same yeast species. In the “unified nomenclature,” the other twelve subunits of the RP were designated Rpn subunits, for regulatory particle non-ATPase proteins (55). The two largest Rpn subunits, Rpn1p and Rpn2p, are part of the Base domain, along with the six Rpt subunits (72). The Rpn10 subunit, which was first identified as a ubiquitin-binding subunit (44, 48), connects the eight-subunit Base with the Lid domain composed of the other twelve Rpn subunits (72). This was shown when a modified proteasome containing only the 20S core and the Base domain was purified (72) from a yeast *Rpn10* mutant. This proteasome could both unfold and degrade proteins; however, ubiquitinated proteins were not degraded (72). The two non-ATPases of the Base, Rpn1 and Rpn2, are predicted to form a hydrophobic surface which binds partially unfolded substrates in a non-specific manner (269). It has been proposed that the subunits in the Base sub-domain of the RP of the 26S proteasome may be able to discriminate between folded and unfolded states of the protein. In addition, depending on interacting factors (*e.g.*, the 20S proteasome, other subunits in the RP), the Base can target the substrate either for refolding or for unfolding and degradation (24, 72).

AAA-type ATPases

The AAA (ATPases associated with a variety of cellular activities)-type ATPases are a subgroup of the Walker-type NTPase family (reviewed in 37, 170, 186). Walker-type NTPases all have a well-conserved region which contains both the Walker A (GPPGXGKT) and Walker B (DEID box) motifs (271). The Walker A motif binds the β and γ phosphates of nucleotide triphosphates (208), and the Walker B motif is involved in

Mg²⁺-coordination and ATP hydrolysis (27). Members of the AAA family have an additional highly conserved second region of homology (SRH) found C-terminal to the Walker A and B motifs (240). The SRH region has been proposed to have a catalytic role in ATP hydrolysis based on studies of FtsH proteins with point mutations in the SRH. These SRH mutants were able to bind but not hydrolyze ATP (111). The AAA proteins are all Mg²⁺-dependent ATPases which are inhibited by n-ethylmaleimide (NEM), a sulfhydryl alkylating agent which is presumed to bind to a well-conserved cysteine residue in the C-terminal region of the SRH domain (294). All members of this family have very well conserved C-terminal domains but very little if any N-terminal identity. These AAA perform a wide variety of roles in the cell, ranging from organelle biogenesis to proteolysis. A number of these proteins are involved in protein folding and unfolding as well as in the disassembly of protein complexes. The AAA proteins can be divided into those with one AAA module (Type I), and those with two (Type II). The Rpt members of the RP, ClpX, ClpQ (HslU), and FtsH (HflB) proteins are all Type I AAA-ATPases, while ClpA, ClpB, ClpC (219), Hsp104, NSF (241), Cdc48 (57), and Lon are Type II ATPases. Notably, for several of the Type II AAA proteins, one or both of the AAA domains are necessary for oligomerization of the ring complex, and ATP binding (but not hydrolysis) is often necessary for complex formation (89, 225, 297).

AAA-ATPases in the RP of the 26S proteasome. As mentioned above, the six ATPase members (Rpt1p to Rpt6p) of the 19S cap of the 26S proteasome are all Type I AAA proteins. Interestingly, the AAA subunits of the RP are much more similar between species than any of the Rpn (non-ATPase) subunits of the RP. In addition, the Rpt subunits have higher identity to homologs from other species than to paralogs within the

same organism, indicating that these subunits most likely serve non-redundant specific functions within the cell. For example, Rpt2p/S4 in human and yeast is 74% identical, while Rpt2p and Rpt8p within *S. cerevisiae* are only 49% identical at the amino acid level. Likewise, the *M. jannaschii* Rpt2p homolog (PAN) is 49% identical to human Rpt2p at the amino acid level, while the identity of Rpt2p and Rpt8p paralogs within *M. jannaschii* is only 38%. In yeast, substitutions of the conserved ATP-binding residue lysine in the Walker A motif were lethal in four out of the six proteasomal ATPase (Rpt) mutants (74), which is a further indication that these six proteins serve non-redundant functions in the cell (203). Based on the low sequence identity of the N-terminal regions of the AAA subunits of the RP, it has been suggested that this region is responsible for substrate selection and/or binding to specific Rpn subunits in the Lid sub-domain of the RP.

Other AAA-ATPases. One well-characterized AAA protein involved in eukaryotic organelle biogenesis is Cdc48 (57), which mediates ATP-dependent fusion of ER membranes. This protein forms homohexamers in the cytosol and was first identified by complementation of mutants which arrested with an undivided nucleus during mitosis (57). One other homolog of Cdc48 is p97/VCP (valosin-containing protein), which also forms a homohexameric ring and is involved in membrane fusions in the Golgi apparatus (189). Other well-characterized Type II AAA proteins in eukaryotes include NSF (NEM-sensitive fusion protein) and its yeast homolog Sec18p, which are proteins involved in mediating vesicle fusion events and transport between the ER and the Golgi apparatus in the secretory pathway. The NSF is required for protein transport from the ER to the Golgi complex, for fusion between endosomal vesicles, and for transport of newly synthesized

proteins to the plasma membrane (165). The NSF forms cytosolic homohexamers as well as larger membrane-associated complexes with the aid of soluble NSF attachment proteins (SNAPs) and SNAP receptors, or SNAREs (165). The NSF disassembles the SNAP/SNARE complexes in an ATP-dependent reaction, which allows another cycle of membrane fusion to occur (89). The N-terminal domain of NSF has been shown to be the substrate-binding domain and is not required for hexamerization of the complex (241, 281). The first AAA module of NSF, D1, has high ATPase activity which is required for a conformational change to occur which mediates disassembly of SNARE complexes (89). The second module of NSF, D2, has very high affinity for ATP but very low ATPase activity (89). The binding of ATP or a non-hydrolyzable analog to the D2 module is required for proper oligomerization of the hexameric complex in the cytosol, and when D2 is expressed alone, a hexameric complex still forms (241, 281). The crystal structure of the D2 hexamerization domain has been solved to 1.75Å resolution in complex with both ATP and the nucleotide analog AMP-PNP (143, 297). The NSF-D2 was found to form a regular hexamer of 97Å in diameter with the nucleotide (ATP or AMP-PNP) bound between neighboring subunits of the ring (143, 297).

A member of the AAA family has also been isolated from the gram-positive actinomycete *Rhodococcus erythropolis* (289). When this protein was synthesized in *E. coli*, the subunits formed a six-membered ring-shaped complex with ATPase activity, which the authors termed ARC (AAA-ATPase Ring-shaped Complex). The *arc* gene was found 1500 bp upstream of an operon containing the 20S proteasome genes. However, interaction of ARC with the 20S proteasome was not reported.

Chaperone-like activity of AAA proteins. Ironically, a study of the mechanisms of proteolysis in the cell must also include a study of the mechanisms of protein folding. On the surface, protein folding and assembly would appear to be completely different processes from protein degradation. However, it has become evident that some of the same proteins involved in degradation of protein substrates can also act as chaperones to aid in the folding of proteins, at least *in vitro* (reviewed in 239). In addition, studying the chaperone and unfoldase activities of ATPase components of ATP-dependent proteases independent of the proteolytic components can aid in understanding the mechanism by which protein substrates are unfolded and translocated into the proteasome and other proteases. The AAA proteins which are components of ATP-dependent proteases have been demonstrated to have the ability to serve as chaperones independent of the proteolytic component (reviewed in 80, 239, 300). Both ClpA and ClpX have been observed to prevent aggregation of substrate proteins *in vitro* (145, 178, 278, 284). The ClpB protein (290) as well as its yeast homologue Hsp104 (184) can also serve as molecular chaperones (34, 185), but neither one of these AAA proteins has been reported to associate with a proteolytic component (*e.g.*, ClpP, HslV). There is some evidence that FtsH may also function as a chaperone (210). The N-terminal AAA domain of the membrane-bound Yme1 protease (FtsH homolog) from yeast can operate as a chaperone to bind unfolded protein substrates and prevent their aggregation even when the C-terminal protease domain is either inactivated or removed (144). The Base component of the 26S proteasome which contains the six Rpt subunits also has ATP-dependent chaperone activity independent of both ubiquitin and the 20S proteasome (24). In one study, non-ubiquitinated citrate synthase could be efficiently refolded by the Base

component, and when the 20S proteasome was included in the assay, citrate synthase was still refolded but not very efficiently degraded (24). Taken together, these results lead to the proposal that AAA-ATPases can behave as relatively non-specific “unfoldases” to remodel, disassemble, and unfold proteins; however, the ultimate fate of a substrate protein rests with the protein (*e.g.*, the 20S proteasome) which interacts with the AAA protein. In addition, the final destiny of a substrate protein for either degradation or refolding can depend on the balance between affinity of the substrate for the AAA-ATPase as well as the stability of the interaction between the AAA component and the proteolytic component (reviewed in 80, 239, 300).

The RP of the proteasome has also been shown to have a chaperone-like role in the yeast nucleotide excision repair (NER) pathway for UV-damaged DNA (205). The 26S complex binds to Rad23, a protein which is required for NER and which also contains a ubiquitin-like domain. Removal of the ubiquitin-like domain from Rad23 or inhibition of RP complex AAA-ATPases resulted in UV-sensitivity. However, inhibition of the proteolytic activity of the 20S proteasome did not result in sensitivity, indicating that the RP complex has a chaperone-like role independent of the proteolytic activity of the 20S proteasome (205). It has been proposed that the 19S complex may bind to the Rad23/excised DNA complex and disassemble the complex to allow for recycling of Rad23 (and other proteins involved in NER). This pathway bears some similarity to the NSF protein disassembly pathway, which disassembles SNARE/SNAP complexes to enable further rounds of membrane fusion and protein trafficking in the ER to take place. Notably, this study provided an *in vivo* example of the Base domain of the RP behaving

as a chaperone and provided further evidence that the chaperone activity which had been previously observed *in vitro* may also have physiological relevance.

ATP-dependent Proteases in *E. coli*

Energy-dependent proteolysis is also important in the maintenance of the eubacterial cell. Several two-component ATP-dependent proteolytic systems have been identified in *E. coli*, including ClpAP, ClpXP, and ClpYQ/HslUV (reviewed in 78, 193). In analogy to the eukaryotic 26S proteasome, each of these proteases contains a proteolytic component (ClpP or ClpQ/HslV) and an ATPase regulatory component (ClpA, ClpX, or ClpY) (209, 277).

The ClpAP and ClpXP proteases. The ClpP protein is a serine protease which forms a double-ringed complex with seven-fold symmetry (224). The ClpP protease can associate with either of the hexameric ATPase components ClpA or ClpX (82). While the ClpAP and ClpXP proteases appear to be quite similar to the proteasome in overall structure based on electron microscopy (117), the crystal structure of ClpP revealed a unique "hatchet-like" fold of its subunits (273, 274), unlike the Ntn-hydrolase β -sandwich fold of the 20S proteasome subunits. The AAA components ClpA or ClpX provide substrate specificity to the ClpP proteolytic complex. The ClpA ATPase can either target the substrate RepA for degradation by ClpP or act alone as a chaperone to disassemble RepA dimers into monomers (284). The ClpX ATPase acts to disassemble and/or degrade bacteriophage initiation protein λ O aggregates (278, 288) and MuA-DNA tetramers (138). Assembly of both ClpA and ClpX with ClpP requires ATP binding but not hydrolysis (100, 230). When wild-type ClpA is purified from *E. coli*, it forms a homo-hexameric complex composed of 84-kDa subunits (104). However, when ClpA

was overproduced using a plasmid in *E. coli*, both full-length (ClpA84) and an N-terminally truncated form (ClpA65) were produced due to an internal translation initiation site, and a hexameric complex containing both proteins was purified (217). When a mutant strain was constructed to produce the truncated ClpA65 protein alone, a hexameric complex still formed. However, this complex had only 5% of the ATPase activity of a ClpA84 complex and in addition could not support ATP-dependent degradation of casein. A complex formed of the full-length ClpA84 alone could support 2.5-fold higher ATP-dependent casein degradation than the ClpA84/65 mixed complex (217).

The HslUV (ClpYQ) protease. The proteolytic component in *E. coli* with the most similarity to the proteasome is the HslUV (ClpYQ) complex (201), which has been called both the “prokaryotic proteasome” and a “hybrid” of Clp and the proteasome (19). The subunits of the proteolytically active HslV (ClpY) component have some sequence identity (~20%) to the β subunits of archaeal and eukaryal proteasomes (201, 202, 295), with the same β -sandwich crystallographic fold (19) and N-terminal threonine active site (225). Interestingly, there has been one HslV homolog (with 52% identity to the *E. coli* HslV) identified in *Bacillus subtilis* which uses an N-terminal serine residue as its catalytic nucleophile (110). In contrast to the proteasome, the HslV protease forms a double ring with six-fold rather than seven-fold symmetry, while the AAA component HslU (ClpQ) also forms a six-membered ring (118). There has been no protein identified which serves the function of the α subunits in the 20S proteasome. However, the HslV complex is not proteolytically active unless it has formed a complex with the ATPase component HslU (295), and the opening into the catalytic chamber formed by the ring of

HslV subunits is quite narrow (19Å). The crystal structure of HslV has been solved both alone (19) and in complex with the AAA regulatory component HslU (233). This was the first solution of a crystal structure of both components of a two component ATP-dependent protease while in complex together. The authors confirmed that the complex was enzymatically active by assaying the crystals for peptidase activity and by using small angle X-ray scattering to solve the quaternary structure of HslUV in solution under conditions where the complex was active (233). The HslU component contains two AAA domains, one in the N-terminal region and one in the C-terminal region. Both were quite similar to the D2 ATPase domain of the NSF protein (see above) (233). The HslU complex also contains a novel feature between the N- and C-terminal AAA modules, which has been termed the “intermediate domain.” This domain protrudes outward from the HslU hexamer away from the HslUV complex (233) and most likely has a role in substrate selection and binding. When the HslU and HslV complexes bind to one another in the presence of ATP, several conformational changes occur. The N- and C-terminal AAA modules of the HslU complex form a cleft which is filled by the N-terminal apical helical region of the HslV subunits. In addition, the C-terminal 14 residues of each HslU subunit form a well-defined α helix which extends to bind between subunits of the HslV protease to form the active HslUV complex (233).

The Lon protease. Unlike the 26S proteasome and the Clp family of proteins, the Lon and FtsH/HflB proteases contain the ATPase and proteolytic components on a single subunit (35). Lon is a serine protease which was first purified from *E. coli* as a complex of four identical subunits(35) but has recently been purified from yeast mitochondria as a seven-membered ring complex (237). Each subunit contains an ATPase domain near the

middle of the sequence and an active-site serine near the C-terminus of the polypeptide (reviewed in 239). Interestingly, the ATPase and protease domains of Lon were synthesized separately in the yeast mitochondria and were found to still interact with one another to perform the same functions as the native protein (263). Substrates of Lon include transcriptional activators, cell division proteins, and mediators of the stress response. The *lon* mutants are deficient in septum formation and FtsZ ring formation (reviewed in 239). The Lon protease is highly conserved in all three domains of life, and yeast *lon* mutants can be complemented by overproduction of the *E. coli* Lon protein (248).

FtsH (HflB) proteases. The FtsH (HflB) protein, a zinc metalloprotease, contains two membrane-spanning domains near the N-terminus, followed by an AAA domain and a Zn^{2+} -dependent proteolytic-active site (HEXXH, where X is any uncharged amino acid) in the cytoplasm (reviewed in 210). Recombinant FtsH assembles into homo-multimeric rings based on electron microscopy (227), and this oligomerization is dependent on the N-terminal membrane-spanning region (5). There is also evidence that the cytoplasmic ATPase domains self-interact and that this association is necessary for degradation of soluble proteins (2). Furthermore, FtsH can form a large multi-protein complex in the membrane with several other proteins, including HflKC and YccA (119, 120). The FtsH protease has been shown to recognize substrates based on their C-terminal sequences (93). These substrates include both the membrane-bound SecY translocase (4) and subunit *a* of the F_0 -ATPase (3) as well as the soluble heat-shock factor σ^{32} (109, 254). Overproduction of FtsH can complement *clpP* mutants in *E. coli* (93), and while *ftsH* knockout mutants are not able to grow, *ftsH* temperature-sensitive mutants can be

partially suppressed by overproduction of the chaperones GroEL/GroES (Hsp60/Hsp10) and Hsp90 (226). These results indicate possible overlapping functions between proteases (ClpP and FtsH) and molecular chaperones.

Substrate Targeting and Degradation

One common theme throughout all three domains of life is that the primary sequence of a protein, especially sequences near the N- and/or C-termini, ultimately determine where a protein will be localized in the cell, how it will be processed to become fully active, and when (and how) it will ultimately be degraded. In addition, if a protein is unstable or partially unfolded, sequences on the protein which would normally be buried (*e.g.*, hydrophobic residues) become exposed to the environment. These can also serve as signals for either chaperones or proteases. In addition, the environment of the cell (*e.g.*, heat-shock conditions, oxidative stress) will affect the thermodynamic stability of cytosolic proteins (192).

Ubiquitination. In eukaryotic cells, substrate recognition in the cytosol is usually at the level of ubiquitin conjugation, where numerous proteins are involved in targeting substrates to the 26S proteasome (reviewed in 36, 94, 95). Ubiquitin is a 76 amino-acid protein which, when conjugated to a substrate protein, targets it for degradation by the 26S proteasome complex. First, ubiquitin is activated and linked at its C-terminus to a ubiquitin-activating enzyme (E1) in an energy-dependent reaction via a ubiquitin-AMP intermediate. The ubiquitin is then transferred to a thiol group on a ubiquitin-conjugating enzyme (E2). The E2 enzyme will then bind to a ubiquitin-protein ligase (E3), which is directly bound to a specific substrate protein. The carboxy-terminus of ubiquitin will then be directly linked to the ϵ -amino group of a lysine residue within the target. After the first

ubiquitin is attached, further rounds of ubiquitination can then occur by the same mechanism, with the incoming ubiquitin being attached via its carboxyl group to (primarily) lysine 48 of the previously-attached ubiquitin molecule. A fourth ubiquitination conjugation factor, E4, has been shown to be essential in recognition of ubiquitinated proteins and in mediating further rounds of ubiquitination (133). After a substrate is ubiquitinated with a minimum of four ubiquitin molecules (252), it is then recognized by subunits in the Lid domain of the 19S cap of the 26S proteasome, such as Rpn10 (44, 264). The ubiquitin chains are then removed by deubiquitinating (DUB) enzymes (such as Doa4/Ubp4) which are also in the Lid of the RP (182), and isopeptidase T/Ubp14 (285) before the protein substrate is fed into the proteolytic core of the complex. Ubiquitin is a very stable protein that is removed intact from the substrate protein and can then re-enter the ubiquitination cycle. While most organisms which have been analyzed contain only one or two E1 ubiquitin-activating enzymes, they contain numerous E2, E3, and DUB enzymes (reviewed in 36, 94, 95). Unlike E2 enzymes, E3 proteins which have been studied have no sequence similarity to one another, even within the same species, and most of them have been identified by binding assays of specific targets in the proteolytic system. Thus, there may be a number of E3 enzymes which are present in a given organism but which have not yet been identified. It is likely that these multiple E2 and E3 enzymes acting in cooperation with one another and with other proteins such as molecular chaperones provide both specificity as well as the ability to target a wide variety of substrate proteins.

Sequences within the substrate protein which will render it more susceptible to recognition by ubiquitination enzymes have also been identified. The "N-end rule"

(reviewed in 265, 266) states that certain residues of a protein at the N-terminus will make it more susceptible to degradation. Replacement of those residues with other amino acids will protect the protein from degradation. In addition, the placement of lysine residues with regards to the N-terminal residue can also play a role in whether or not the protein will be ubiquitinated. Certain residues on the C terminus have also been shown to render a protein more susceptible to degradation (115, 147). Phosphorylation is an additional mechanism which can render a substrate either more or less accessible to ubiquitination proteins (32, 33).

Other eukaryal degradation signals. While ubiquitination is understood to be the primary signal for substrate recognition in eukaryotes, ubiquitination is not necessary for all proteins to be targeted to the proteasome, including ornithine decarboxylase (148) and cyclin-dependent kinase inhibitor p21 (p21^{Cip1}) (221). The exact nature of the signaling involved in these and other substrates is becoming more clear, as there is evidence that PEST sequences on the substrates themselves may be involved (reviewed in 195, 268). These PEST sequences are regions enriched in proline, glutamate, serine, and threonine residues, and uninterrupted by positively charged residues. They are frequently found in rapidly degraded and highly regulated proteins in the eukaryotic cell and have been recognized as degradation signals (200). However, the mechanism for recognition of PEST sequences by the proteolytic machinery is not very well understood. For ornithine decarboxylase (ODC), the PEST sequence is found near the C-terminus and, in combination with a second protein called antizyme, helps to target ODC to the 26S proteasome (69, 70, 148). Antizyme can directly target ODC to the 26S proteasome, bypassing the ubiquitin system. Similar to ubiquitin, antizyme is released intact from the

26S complex, allowing one antizyme molecule to participate in multiple rounds of ODC degradation (149). Cyclin-dependent kinase, on the other hand, is targeted to the proteasome only when the tetrameric complex, composed of two regulatory (R) and two catalytic (C) subunits, has dissociated into monomers. Both R and C subunits contain PEST sequences, but these sequences appear to be buried in the tetramer and are only exposed after cAMP-dependent dissociation of the complex (21). In some cases, PEST sequences can also act in cooperation with the ubiquitin system. The transcriptional inhibitor I κ B α , which contains a C-terminal PEST sequence, is rapidly degraded when dissociated from the transcriptional activator NF- κ B, and deletion of the PEST-containing C-terminal residues stabilizes the I κ B α protein (32). In addition, there are certain cytokines which can target I κ B α for degradation while still in association with NF- κ B (33). In this case, PEST sequences act prior to ubiquitination by serving as a signal in the phosphorylation of two specific serine residues in the I κ B α protein (33). The phosphorylated I κ B α protein is then recognized by ubiquitination enzymes or other members of the proteolytic pathway (195). Interestingly, while some proteins do not appear to require ubiquitination for degradation, some proteins can be labeled with ubiquitin and still not be degraded by the 26S proteasome (105, 129). So, while much progress has been made in understanding the mechanisms of substrate labeling and degradation in eukaryotes, there are still many questions to be answered and much to be learned about this complex process.

Eubacterial degradation signals. While the ubiquitin system is conserved in all eukaryotes examined, neither ubiquitin nor the enzymes of the ubiquitination pathway have been found in any archaeal or eubacterial cells. No gene with significant identity to

ubiquitin has been discovered in any of the prokaryotic or archaeal genomes which have been completely sequenced. In prokaryotic organisms, the model is that proteins are generally directly targeted to the ATPase subunits or domains of the protease by sequences inherent in the proteins themselves.

Proteins which recognize and degrade N-end rule substrates have also been identified in eubacteria (253). Based on genetic screening of *E. coli* mutants in which N-end rule substrates were stabilized, ClpAP (but not ClpXP) has been implicated in this pathway (228). In contrast to eukaryotes, an intermediary protein that directly recognizes the N-terminal residue of N-end rule substrates and targets the substrate to ClpA has not been identified in any prokaryotes. Instead, it has been proposed that substrate recognition may be at the level of the ATPase component of the protease itself (reviewed in 94). Therefore, these ATPases are likely to not only unfold and feed substrates into the proteolytic core (similar to the Rpt subunits of the RP), but they may also directly recognize substrate proteins (similar to the Rpn ubiquitin-recognition subunits of the RP).

Another mechanism for targeting proteins for degradation in eubacteria is the SsrA-tagging system. The 10SaRNA transcript of the *ssrA* gene can add a "degradation sequence" to an RNA transcript without a stop codon or to one that is stalled on the ribosome. The *ssrA* transcript is a 363nt RNA unit which has been termed "tmRNA" because it serves the function of both tRNA and mRNA (116, 256). The tRNA portion of the tmRNA is charged with alanine and can bind to the P site of a stalled ribosome. Translation then switches to the mRNA-like portion of the tmRNA that encodes a 10-residue peptide tag. Thus, the final protein has the sequence AANDENYALAA added co-translationally at the C-terminus. This peptide sequence targets the protein for

degradation by cellular proteases, such as ClpAP, ClpXP, or FtsH (79, 93). Interestingly, ClpAP and ClpXP have been observed to recognize both the N- and C-termini of substrates (77, 138), indicating that one regulatory ATPase can recognize several different types of recognition signals. In addition, other proteins have been shown to be involved in enhancing the specificity of the ATPase component for a tagged substrate, including SspB. The SspB protein is a ribosome-associated protein which binds specifically to SsrA-tagged proteins, targeting them specifically to the ClpX ATPase (146). Even though both ClpA and ClpX can recognize SsrA-tagged substrates *in vitro*, there is evidence that ClpX recognizes the majority of these proteins *in vivo* (79). The *sspB* gene is part of an operon induced by starvation, which is a time of stress when more polypeptide chains are stalled on the ribosome. The upregulation of *sspB* allows for more specific control of which proteins are targeted for proteolysis than there would be if only ClpX itself was upregulated. Specificity factors, such as SspB, help to explain how a single ATPase can recognize a diverse number of substrates and also how different cellular environmental conditions will cause specific proteins to be targeted for unfolding and degradation.

Substrate unfolding and degradation. For many of the energy-dependent protease components, including the 26S proteasome, ClpAP, and ClpXP there is an asymmetrical mismatch between the ATPase component (with a six-membered ring) and the protease component (with a seven-membered ring). It is thought that ATP hydrolysis by the regulatory component causes rotation of the ATPase, allowing the substrate to be threaded into the proteolytic core of the complex (18, 109). Degradation of protein substrates by either ClpAP or ClpXP is a highly processive process (251); that is, a

substrate is completely degraded and released before another substrate molecule is taken up by the protease. Studies indicate that protein degradation by the archaeal 20S proteasome from *T. acidophilum* is also very processive (6).

While degradation by the eukaryotic 20S proteasome has also been demonstrated to be very processive (126), there are a few notable exceptions. For example, the transcriptional activator NF- κ B is a heterodimeric protein consisting of a p65 and a p50 subunit. The p50 subunit is synthesized as a 105-kDa precursor (p105), and the 26S proteasome is responsible for degradation of the C-terminus of p105 to yield the mature p50 protein (180). Interestingly, the C termini of p105 and of I κ B α are very similar. Both contain PEST sequences which lead to phosphorylation of the protein at specific serine residues (180). However, the mechanism by which the proteasome endoproteolytically cleaves p105 into two fragments and then specifically degrades only the C-terminal region is unknown.

Other Proteases

Other proteases have been isolated which can partially compensate for loss of proteasome function in mammalian cells including the subtilisin-like serine protease tripeptidyl peptidase II (TPPII) (64). The ability to compensate for proteasome function was first observed in an outgrowth of mouse lymphoma cells which had adapted to proteasome vinyl sulphone inhibitors (71). The TPPII protein, which had been previously isolated and characterized (12), was later shown to be responsible for this adaptation (64). The TPPII protein forms a rod-shaped structure when observed using electron microscopy and is somewhat larger than the 26S proteasome with dimensions of ~50nm by 17 nm (64). It is an exopeptidase which requires the N-terminus of the protein

substrate to be exposed, and it releases multiples of three residues after cleavage (12). While there is little sequence or structural similarity, it has been hypothesized that this protease may be the functional mammalian homolog of the archaeal tricorn protease (64); that is, it may further degrade oligopeptides after release from the proteasome (discussed below). It also has some sequence identity to the archaeal serine protease pyrolysin (discussed below), which was first isolated from the hyperthermophilic archaeon *Pyrococcus furiosus* (270). However, pyrolysin is an endoprotease rather than an exoprotease.

Archaeal Proteases and Chaperones

The complete genomic sequences of nine euryarchaea have been published: *M. jannaschi* (26), a hyperthermophilic methanogen with an optimal growth temperature of 82° C, originally isolated from a deep-sea hydrothermal vent (108); *Archaeoglobus fulgidus* (128); *Methanobacterium thermoautotrophicum* ΔH (231); *Pyrococcus abyssi* (92); *Pyrococcus horikoshii* OT3 (113); *Thermoplasma acidophilum* (113, 204); *Halobacterium* sp. NRC-1 (171); and most recently *Thermoplasma volcanium* GSS1 (114). The genomes of crenarchaea *Aeropyrum pernix* K1 (112) and *Sulfolobus solfataricus* (220) have also been sequenced. In addition, the sequence of *M. maripaludis* strain LL has recently been completed (J. Leigh, personal communication).

Archaeal proteasomes. Each archaeal genome has ORFs encoding α and β subunits of the 20S proteasome, which are unlinked on the chromosome (See Table 1). Based on genome sequencing, both *Pyrococcus* species as well as *Aeropyrum pernix* have two β subunit paralogs. *Haloferax volcanii* is the only organism at this time which has been shown to have two α subunit paralogs (286); even the completely sequenced

TABLE 1. Proteases and proteolytic components found in the Archaea based on genome sequences.

Protein	PsmA ¹	PsmB ¹	PAN ² (Rpl2p)	Lon/Lon2	Tricorn ³ F1, F2, F3	HtpX	PfpI ⁴	Hsp100 Clp
COG Number	0638	0638	1222	1067	0596, 0308	0501	0693	0542
Organism								
<i>A. fulgidus</i> (128)	AF0481	AF0490	AF1976	AF0364; AF2060	---	AF0235	AF1281	---
<i>A. pernix</i> (112)	APE1449	APE521; APE507	APE2012	---	---	APE1045	APE319	---
<i>Halobacterium</i> sp. NRC-1 (171)	AAG18783 (annotated-PsmB)	AAG19322 (annotated-PsmA)	AAG19042 AAG20171	AAG18884	---	HSP4	AAG20607	---
<i>M. jannaschii</i> (26)	MJ0591	MJ1237	MJ1176 (41)	MJ1417; MJ1489	---	MJ1682	MJ0967	---
<i>M. thermoautotrophicum</i> (85)	MTH686	MTH1202	MTH0728	MTH785; MTH892	---	MTH569	---	MTH284
<i>P. abyssi</i> (92)	PAB0417	PAB1867; PAB2199	PAB2233	PAB1313	---	PAB758 PAB1974	PAB311	---
<i>P. horikoshii</i> (113)	PH1553	PH1402; PH245	PH201	PH0452	---	PH1256; PH1588	PH1704	---
<i>T. acidophilum</i> (204)	TA1288 (13)	TA612 (13)	---	TA1081; TA0098	TA0830; TA0301; TA0815	TA1490 TA861	TA0465	---
<i>T. volcanium</i> (114)	TVG0317752	TVG0656650	---	---	TVG 0817891	TVG 0940548	---	---

¹20S proteasomes have also been identified in the following organisms: *M. thermophila* (157), Accession numbers: α , T48878; β , T48879; *H. volcanii* (286), Accession numbers: α , T48678; β , T48677; *P. furiosus* (13), genome sequences not available; *Thermococcus* sp., Accession numbers: α , T43887; *S. solfataricus*, Accession numbers: α , CAB57565; β , CAB23181.

²PAN homologs have also been identified in the following organisms: *M. maripaludis* (Wilson and Maupin-Furlow, this work); *H. volcanii*: contains two paralogs (Maupin-Furlow, unpublished).

³Tricorn protease homologs have also been identified in *S. solfataricus*, Accession numbers S73091; p95871; as well as the aminopeptidase interacting factors, Accession numbers CAA69432; AAC63218 (214).

⁴PfpI homologs have been identified in *P. furiosus* (88), Accession number: AAB04694; *S. solfataricus*, Accession numbers CAC23044, AAG20607. ---, not identified.

Halobacterium sp. NRC-1 appears to have only one α subunit (171). All Archaea examined, with the exception of the euryarchaea *T. acidophilum* (204) and *T. volcanium* (114), also contain at least one AAA-ATPase with high identity to regulatory subunit Rpt2p (human S4) of the eukaryotic 26S proteasome. The Rpt2p homolog of *M. jannaschii* has been purified and characterized (303). The authors called the protein PAN for Proteasome Activating Nucleotidase. This protein could activate the ATP-dependent protein hydrolyzing activity of both *T. acidophilum* and *M. thermophila* 20S proteasomes. It is notable that *H. volcanii*, the only organism that has been shown to contain two α subunits of the 20S proteasome, also has two PAN paralogs (J. A. Maupin-Furlow, unpublished results).

One especially intriguing difference between eukaryotes and the archaea is that, even though archaea contain the core components of the 26S system (the 20S proteasome and at least one homolog of the ATPase subunits of the 19S cap), no homolog of the subunits involved in recognition and removal of ubiquitin has been identified. In addition, none of the ubiquitin activation, conjugation, or ligation (E1, E2, and E3) enzymes have been identified. Furthermore, a gene encoding ubiquitin itself is not predicted to be present in any of the archaeal genomes that have been sequenced. However, it is likely that there is an as yet unidentified system that specifically targets substrates to ATP-dependent proteases in Archaea. For example, a small protein (similar to ubiquitin) or peptide (similar to SsrA) tags larger substrates and targets them for degradation by the ATP-dependent proteasome.

Lon and FtsH protease homologs. The Lon protease is highly conserved throughout all three domains of life and is predicted to be present in all archaeal genomes

which have been sequenced. The proteolytic serine residue active site as well as residues involved in ATPase activity are well-conserved based on multiple-sequence alignments. Interestingly, the archaeal Lon homologs all have putative N-terminal membrane-spanning regions, unlike the eubacterial and eukaryal homologs. This suggests that they are associated with the membrane in the archaeal cell (158). The function of Lon in the archaeal cytosol is as yet unknown. An ORF has also been identified in archaea which was originally identified as the Rpt6p (human S8) ATPase regulatory subunit of the 26S proteasome but which also has some identity to the FtsH protease. The FtsH homolog in the archaea does not appear to have the Zn^{2+} -dependent active site or a membrane-spanning region. So, in contrast to archaeal Lon, the FtsH homolog is most likely soluble. It has been proposed that the membrane-bound archaeal Lon protease may substitute for FtsH function in the Archaea (204), and that the FtsH/Rpt6p homolog is not a protease but rather a molecular unfoldase or chaperone involved in protein degradation or folding (158).

The tricorn protease. One other novel protease, the tricorn protease, has thus far been studied only in the thermophilic archaeon *T. acidophilum* (244). The tricorn protease is a large macromolecular assembly consisting of 121-kDa monomers which associate into hexamers, 20 of which can then finally assemble into a large icosahedral capsid of 14.6 MDa (242, 272). The capsid serves as an “organizing center” for three interacting factors, including a proline iminopeptidase (245) and two amino-peptidases (242). The tricorn protease core assembly can degrade oligopeptides of 6-12 residues processively from the N-terminus into 2-4 residue peptides, and the interacting factors further degrade these peptides into single amino acids (242). Furthermore, the tricorn

protease is an exopeptidase which requires the N-terminus of the substrate to be exposed. Based on the size of the substrates degraded, it is thought that the tricorn protease operates downstream from the proteasome in the *T. acidophilum* cytosol to further degrade peptides into amino acids. In addition, the entire proteolytic pathway has been reconstituted *in vitro* using recombinant 20S proteasome, tricorn protease, and the interacting factors (242). Based on sequence analysis, *Sulfolobus solfataricus* and *T. volcanium* both contain a tricorn protease homolog as well as the interacting factors. However, the Tricorn protease core has no sequence identity to any ORFs from other sequenced archaeal genomes, though both the proline iminopeptidase and the aminopeptidase components share sequence identity to other zinc-dependent amino-peptidases from all three domains of life (242).

Other Archaeal proteases. The PfpI protein is a novel archaeal intracellular protease which was first purified from *P. furiosus* (88) and which also appears to be conserved in all other completely sequenced archaea with the exception of *M. thermoautotrophicum*. In addition, the Clp system does not appear to be in any of the archaea which have been sequenced, with the exception of one gene from *M. thermoautotrophicum* which has some sequence similarity to Clp (Hsp100) ATPase genes from eubacterial organisms. This protein may function solely as a chaperone without a protease component (similar to ClpB) or may perhaps even cooperate with other chaperone systems, as ClpB does with the DnaK/DnaJ/GrpE proteins in *E. coli* (179). However, in contrast to other Clp proteins, the *M. thermoautotrophicum* Clp homolog appears to be a Type I AAA-ATPase with only one AAA module. None of the

other Archaea that have been completely sequenced appear to contain this protein, including the mesophilic *M. maripaludis*.

Extracellular archaeal proteases. Many archaeal organisms, especially hyperthermophiles and halophiles, grow very well on protein substrates (270). Extracellular proteases are necessary for the organism to degrade and utilize these protein substrates. Numerous extracellular proteases have been isolated from hyperthermophiles, including pyrolysin from *P. furiosus* (270), a subtilisin-like serine protease with high identity to eukaryal tripeptidyl peptidases II (TPPII). However, in contrast to eukaryal TPPII, pyrolysin is an endopeptidase rather than an exopeptidase (270).

Archaeal chaperones. Archaeal organisms reside in a wide variety of environments; therefore, their mechanisms for dealing with stress are of special interest (reviewed in 154, 158). Archaeal cells do not appear to have the eubacterial GroEL(Hsp60)/GroES(Hsp10) chaperone system, based on genome sequence analyses (Table 2). The most well-conserved chaperonin among the Archaea is the thermosome (Hsp60), which has high sequence identity to the TriC(CCT) eukaryotic chaperonin (reviewed in 86, 130). In eukaryotic cells, the TriC chaperonin is a double ring complex consisting of eight different subunits, each in a defined position within the ring (150). The archaeal thermosome, which was first purified from *Pyrodictum occultum* cells exposed to heat shock (190), is also composed of either eight or nine-membered double rings. In contrast to the eukaryotic counterpart, the thermosome complexes purified from *P. occultum*, *Sulfolobus shibatae*, (255), *Thermococcus* sp. KS-1 (296), and *T. acidophilum* (172) consist of only two types of subunits, alternating α and β subunits

TABLE 2. Chaperones and heat-shock proteins found in the Archaea based on genome sequencing.

Protein	Hsp20 sHSP	Hsp60 Thermosome ¹	Hsp60 GimC	Hsp70 DnaK ²	Hsp40 DnaJ	Hsp23 GrpE	Cdc48 VAT ³	Cdc48 VAT2	FtsH Rtp6p/FtsH
COG Number	0071	0459	1730	0443	0484	0576	0464	0464	1223
Organism									
<i>A. fulgidus</i> (128)	AF1296	AF2238; AF1451	AF2063	---	---	---	AF2098	AF1297	AF1285
<i>A. pernix</i> (112)	APE1950	APE2072; APE0907	---	---	---	---	APE2474	APE1367; APE0960	APE2474
<i>Halobacterium</i> sp. NRC-1 (171)	AAG18726; AAG20020; AAG20865	AAG20351; AAG20240	AAG20539	AAG19026	AAG19025	AAG19028	AAG19919	AAG19776; AAG20864	AAG19919
<i>M. jannaschii</i> (26)	MJ0285 (122)	MJ0999 (136)	MJ0952; MJ0648	---	---	---	MJ1156	---	MJ1494
<i>M. thermocauto-</i> <i>trophicum</i> (85)	MTH0859; MTH1366	MTH0794; MTH0218	MTH1609	MTH1290	MTH1291	MTH1289	MTH1639	---	MTH1011
<i>P. abyssus</i> (92)	PAB2072	PAB2341	PAB1006	---	---	---	PAB1478	PAB2086	PAB2086 PAB1478
<i>P. horikoshii</i> (113)	PH1842	PH0017	PH0527	---	---	---	PH0687	PH1840	PH1840
<i>T. acidophilum</i> (204)	TA0437; TA0471; TA0864	TA0981; TA1276 (172)	TA1076; TA1137	TA1087	TA1088	TA1086	TA0840 (181)	TA1175	TA840
<i>T. volcanum</i> (114)	TVG0791352 TVG1034799	TVG0494666; TVG1181974	---	TVG0471466	TVG0470348	TVG0473313 (annotated- GroE)	---	TVG0373521 (annotated- PAN)	---

¹Thermosomes have been identified in the following organisms: *P. occultum* (191), *P. kodakarensis*, *M. kandleri*, *Thermococcus* sp. Ks-1 (296), *S. acidocaldarius* (SAV), (38), *Pyrococcus* sp. KOD1 (296), and *H. volcanii* (137).

²Hsp70/DnaK homologs have also been identified in *M. maezi* and *M. thermophilus* (99).

³VAT protein homologs have also been identified in the following organisms: *S. acidocaldarius*, Accession number AAA72002; CdcH from *H. salinarum* (54), Accession number CAA56097; *P. kodakarensis*, Accession number BAA87866; and *S. solfataricus*, Accession number CAC22860.

---, not identified.

arranged in an eight- or nine-membered ring. Sequences from *H. volcanii* (137), *A. fulgidus* (85), *M. thermoautotrophicum* (112), *T. volcanium*, (114), and *A. pernix* (112) indicate that these organisms also contain two thermosome subunits. On the other hand, some of the hyperthermophilic Archaea including *M. jannaschi* (136) and *Pyrococcus* sp. (293) have only one gene with high identity to other archaeal thermosome subunit genes (Table 2). The eukaryal TriC/CCT mediates proper folding of proteins in the eukaryal cytosol, including the cytoskeletal proteins actin (61) and tubulin (292). The thermosome most likely serves a similar role in archaeal cells. In addition, the thermosome appears to be the predominant heat-shock protein in the cytosol of the hyperthermophilic archaeon *P. occultum* (191). The mechanism for both the archaeal thermosome and the eukaryotic TriC(CCT) does not appear to be dependent on a separate chaperonin component such as GroES. Based on crystal structures of the *T. acidophilum* thermosome (45, 131), it has been proposed that an N-terminal protusion in the substrate-binding apical region of the thermosome serves as the functional equivalent of GroES. These chaperones have thus been named Group II chaperonins to distinguish them from Group I chaperonins such as GroEL which require a separate co-chaperonin for activity. However, there is some evidence that there may be a component which interacts with TriC. This protein, which is called prefoldin (262) or GimC (65), has been shown to be able to bind actin and transfer it to the TriC complex. Prefoldin is also retained on a TriC-ATP-agarose column (262). The GimC/prefoldin complex forms a hexameric ring of six different paralogous subunits in the eukaryal cytosol (262), but it is not known exactly how this would interact with the octameric TriC complex. Homologs of GimC/prefoldin are conserved in the archaea and

eukarya, while they are absent from the eubacteria. Archaea generally have one or two GimC/prefoldin paralogs (Table 2), while eukarya contain six.

The Hsp70 system, consisting of the DnaK protein as well as the cochaperones DnaJ (Hsp40), and GrpE, is very well conserved in both eukaryotes and eubacteria (reviewed in 154). Surprisingly, it is not as well conserved in the archaea. It is also intriguing that the archaeal DnaK/DnaJ/GrpE proteins are more similar at the amino acid level to eubacterial than to eukaryal chaperones. In addition, the archaeal genes are usually found in an operon in the order 5'-*grpE-dnaK-dnaJ*-3', which is identical to the order found in eubacteria; however, the intergenic promoter regions and transcription start sites are eukaryotic-like in nature (reviewed in 154). The DnaK/DnaJ/GrpE system appears to be in place in *Methanosarcina* species (99), *Halobacterium* species, *M. thermoautotrophicum*, and *T. acidophilum*, while it is not present in *Methanococcus* species (including *M. maripaludis*), *Sulfolobus* species, *Pyrococcus* species, or *A. fulgidus*, based on genome sequence analysis (Table 2). It has been proposed that the DnaK/DnaJ/GrpE chaperone system is replaced by another unidentified system in the majority of hyperthermophiles which may be tailored to the special conditions of their environment (114).

The VAT (VCP-like ATPase of *T. acidophilum*) is an archaeal homolog of the eukaryotic Cdc48 and p97/VCP chaperones (181). All sequenced archaeal genomes appear to contain at least one or two homologs of this protein (Table 2). In addition, homologs have been identified and characterized in the partially sequenced genomes of *S. acidocaldarius* (38) and *H. salinarium* (54). When produced in recombinant *E. coli*, VAT from *T. acidophilum* forms single and double homohexameric rings which resemble

Cdc48 and p97 (57, 189) and further associates to form four-ringed structures that resemble the 20S proteasome based on electron microscopy (181, 199). Similarly to NSF and other Type II AAA-ATPases, the N-terminal domain is not necessary for oligomerization of the complex and appears to be quite flexible (199). In addition, AAA domains D1 and D2 of VAT and NSF are quite similar to one another structurally and can be superimposed (199). In contrast to NSF and Cdc48, VAT is stable in the absence of ATP (76). Interestingly, VAT can act to either fold or unfold a protein substrate *in vitro*, depending on the concentration of Mg^{2+} and ATP (76). It has been proposed that this chaperone (and others) may help to compensate for the lack of a PAN homolog in *T. acidophilum* (76).

Small heat shock proteins (sHSPs/Hsp20s) are present in all archaeal genomes sequenced, and a sHSP homolog from *M. jannaschii* has been cloned, purified, and characterized (122). When produced in *E. coli*, the 16.5 kDa protein formed a homooligomeric ~15-20nm spherical structure with a molecular mass of ~400 kDa (122). In addition, *E. coli* cell extracts containing this protein were protected from thermal denaturation and precipitation when heated to 100°C (122). The purified protein also protected both single-chain monellin and citrate synthase from aggregation at 80° and 40° C, respectively (122). The crystal structure of the *M. jannaschii* sHSP was obtained at 2.9Å resolution (121) and revealed that the 24-subunit complex formed a hollow sphere of octahedral symmetry with eight triangular and six square openings. The openings were proposed to be possible inlets/outlets for substrate proteins. It was also hypothesized that substrates could be “trapped” for folding inside the hydrophobic interior of the hollow sphere similar to the GroEL/ES or thermosome complex (121).

As can be seen from the discussion above, while there are many similarities among all archaeal organisms, there are also intriguing differences. These include the lack of a PAN homolog in some organisms, the lack of the DnaK/DnaJ/GrpE chaperone system in hyperthermophiles, the lack of a Pfpl protease homolog (but the presence of a Clp homolog) in *M. thermoautotrophicum*, and the presence of the tricorm protease in two archaeal species. Some attempts have been made to correlate the presence or lack of certain chaperones and other components with the optimal growth temperature of the archaeon (114). However, when halophiles and mesophilic methanococci are also considered, the picture becomes much more complicated. In addition, reconciling the fact that archaeal cells contain both a eukaryotic-like chaperonin system (the thermosome), and, in some cases, a bacterial-like chaperone system (Hsp70/Hsp40/GrpE) has been the subject of much speculation.

Methanogens and Genetic Studies in Methanogens

Performing genetic analyses in methanogenic archaea can be quite challenging, because most of the genetic techniques which have been developed for use in gram-negative eubacteria cannot be applied to these organisms. The strict requirement for anaerobic conditions as well as the fragile S-layer cell envelope of *Methanococcus* sp., which lyses in hypotonic conditions, can make electroporation and CaCl₂-mediated transformation very inefficient. In addition, archaea have a different replication mechanism than eubacteria (29, 50) so broad host-range vectors fail to replicate in the archaea. The RNA polymerase is more complex than the eubacterial counterpart, and some of the transcription factors such as TBP (75) and TFIIB are homologous to the eukaryal proteins. Furthermore, the promoter regions are eukaryotic-like with a TATA

box and initiator element, so eubacterial promoters are not recognized in the archaeal cell. Most antibiotics developed for use in eubacteria fail to inhibit the growth of archaea due to cell wall differences as well as different protein synthesis mechanisms such as ribosome structure. Other techniques such as blue/white selection using the reporter gene *lacZ* and the substrate X-gal must be modified (e.g., by replica plating before screening) because the formation of the blue color produced from the breakdown of X-gal requires oxygen.

Thermophilic organisms present special challenges because cultivation can be time consuming and difficult, especially on solid media; antibiotics as well as the protein which confers antibiotic resistance may be unstable at high temperatures; and unprotected vector DNA may also be unstable at high temperatures (reviewed in 174). Auxotrophic mutant screening is also difficult due to a lack of growth of most thermophilic organisms on defined media in the laboratory (174). Therefore, while some advances have been made in the hyperthermophilic species *Sulfolobus* and *Pyrococcus* (174), little is known about the metabolism or physiology of *M. jannaschii* other than that which can be determined from analysis of the genome sequence (213). In contrast, two mesophilic methanogen species, *Methanococcus voltae* and *M. maripaludis*, have been studied extensively using genetic techniques in recent years. Both of these methanogens have simple protein S-layer cell envelopes, which are easily disrupted for DNA isolation as well as for spheroplast formation. In addition, these methanogens generally grow as non-flagellated single cells, with little clumping, allowing isolation of single cells and colonies (107). The relatively rapid doubling time and mesophilic nature of these

microorganisms have made them popular models for performing genetic analyses in methanogens (reviewed in 258 and 283).

Genetic studies in *M. voltae*. *M. voltae* was isolated from the Waccasassa estuary near Cedar Key, Florida (276) and has a doubling time of 1.2 hours with an optimal growth temperature of 38°C (282). *M. voltae* was the first methanogen in which transformation was successfully performed. In this study, the authors were able to rescue histidine and purine auxotrophs by transformation with wild-type chromosomal DNA (17). In a later study, *M. voltae* was transformed with the integration vectors Mip1 and Mip2 (68). These vectors were derived from plasmid pUC18 (containing an ampicillin resistance marker and an origin of replication for manipulations in *E. coli*). In addition, the vector contained the *pac* (puromycin transacetylase) gene from *Streptomyces alboniger* (139) with promoter and terminator regions from the *mcr* (methyl CoM reductase) operon of *M. voltae*, for puromycin resistance selection in *M. voltae*. The *pac* cassette was flanked by *hisA* gene sequences to promote homologous recombination and integration into the chromosome (68). Histidine auxotrophs could then be isolated and analyzed. In both of these studies, the transformation efficiency was very low (~10 transformants per µg DNA). The reporter genes *uidA* (β-glucuronidase from *E. coli*) and *treA* (trehalose gene from *Bacillus subtilis*) have also been used in *M. voltae* to study the regulation of [NiFe]-hydrogenase genes (14, 16). *M. voltae* contains four sets of hydrogenase genes (87). Two sets (*fru* and *vhu*) contain selenocysteine and are constitutively transcribed. The other two sets (*frc* and *vhc*) contain a cysteine rather than a selenocysteine residue at the homologous position and are transcribed only in the absence of selenium (15). The *frc* and *vhc* genes are divergently transcribed from a

common 453-bp intergenic promoter region. When either the *frc* or the *vhc* gene was replaced with the reporter gene *uidA*, β -glucuronidase activity could be measured only in cells grown in the absence of selenium (14). Furthermore, by placing the intergenic region between *uidA* and *treA* in the same construct, it could be shown directly that these two genes are coordinately regulated (16). Finally, by mutating specific regions of this intergenic promoter region in the above constructs, it was found that these two operons were both under the control of a single negative regulatory element (similar to a eukaryotic silencer) as well as activator regions specific for each of the two hydrogenase operons (173).

One limitation in the use of *M. voltae* as a model organism to study methanogen physiology is that while it can synthesize the amino acids isoleucine and leucine, it has a strict growth requirement in the laboratory for these amino acids. This makes screening for auxotrophic mutants difficult (223). In addition, no plasmid has been found in this organism, which has prevented the development of a self-replicating shuttle vector. Currently, the transformation efficiency is no higher than the frequency of spontaneous puromycin resistant mutants. This makes statistical analyses of transformants and auxotrophic mutants difficult.

Genetic studies in *M. maripaludis*. *M. maripaludis* was isolated from a salt-water marsh and has a doubling time of ~2.3 hours with an optimal growth temperature of 37°C (106). In contrast to *M. voltae*, *M. maripaludis* is a nutritionally versatile facultative autotroph. *M. maripaludis* will assimilate amino acids, acetate, and formate, but the cells can also grow with CO₂ as the sole carbon source. These characteristics have made possible both auxotrophic mutant screening as well as labeling studies to study carbon

metabolism. Before discovery and development of any antibiotic resistance genes in methanogens, acetate auxotrophs were isolated after enrichment using the base analogs 6-azauracil and 8-azahypoxanthine (140). This study provided the first direct genetic evidence that methanogenic Archaea contain the Ljungdahl-Wood pathway for autotrophic CO₂ fixation and acetyl-CoA biosynthesis.

Integrative vectors have been developed for use in *M. maripaludis* (63), based on the original Mip vectors from *M. voltae*, by using a fragment of *M. maripaludis* genomic DNA in place of the *hisA* gene sequences from *M. voltae* to allow for homologous recombination into the chromosome. An integrative expression vector, pWLG14, was also developed by placing a multiple cloning region downstream of the *M. voltae* histone promoter (PhmvA) which allows for constitutive overexpression of a protein of interest after integration into the chromosome (63). A PEG-based transformation method that yields 10⁵ to 10⁷ transformants per µg DNA has also greatly aided genetic studies in this and other methanogens (260). In contrast to *M. voltae*, no spontaneous puromycin resistance mutants are observed in *M. maripaludis*, so analysis of transformation frequency is more direct.

The 8,285-bp cryptic plasmid pURB500, which was isolated from *M. maripaludis* strain C5 (291), has formed the basis for development of an expression shuttle vector for use in this organism. The vector, pWLG30, is 12.7-kbp and was constructed using the integrative expression vector pWLG14 as well as a large portion of pURB500 (259). This self-replicating vector is stably maintained in *M. maripaludis* and supports constitutive expression of heterologous genes. No inducible promoter has been discovered at this time, so control of the timing and level of expression of the protein of interest is not

possible. The expression vector is useful for producing proteins which are oxygen sensitive and/or may require additional archaeal factors for activity. Unfortunately, this vector will not replicate in other closely related *Methanococcus* species, including *M. voltae* (259). In addition, a minimal replication region has not been determined for this plasmid; thus, large portions of pURB500 are required for replication in *M. maripaludis*. The pURB500 vector contains 18 ORFs, most of which seem to be important in maintaining the vector in the methanogen cell (63, 259). Recently, a smaller vector, pWLG40, has been constructed. This plasmid is 10.2-kbp and is maintained in *E. coli* and *M. maripaludis* at a higher copy number than pWLG30 (personal communication, W.Gardner and W. Whitman).

In conclusion, developing a model for protein folding and degradation in archaeal organisms will be a challenging and enlightening task. This will be greatly aided by obtaining the complete genome sequences of some archaeal organisms for which genetic systems have been developed such as *H. volcanii*, *M. maripaludis*, and *M. voltae*. In addition, using some of the genetic techniques which have been developed for use in these organisms should answer some of the questions which have been raised in this review as well as in others.

MATERIALS AND METHODS

Materials

Biochemicals were purchased from Sigma Chemical Co. (St. Louis, Mo.). Other organic and inorganic chemicals were from Fisher Scientific (Atlanta, Ga.) and were analytical grade. Restriction endonucleases and DNA modifying enzymes were from New England Biolabs (Beverly, Mass.) or Promega (Madison, Wis.). Oligonucleotides were from Genemed Synthesis (San Francisco, Calif.) or from Sigma-Genosys (Woodland, Tex.). Positively-charged nylon membranes, digoxigenin-11-dUTP (DIG), and alkaline phosphatase-conjugate antibody raised against DIG were from Roche Biochemicals (Indianapolis, Ind.). Polyvinylidene difluoride (PVDF) and positively-charged nylon membranes were from MicroSeparations (Westborough, Mass.) and Ambion (Austin, Tex.), respectively. The λ BlueStar vector and phage packaging extracts were from Novagen (Madison, Wis.). Morph site-specific plasmid DNA mutagenesis kit was from 5 Prime $\rightarrow$ 3 Prime (Boulder, Colo.). Zero Blunt-TOPO PCR cloning kit and the pCR-TOP10 vector were from Invitrogen (Carlsbad, Calif.). Goat anti-rabbit IgG alkaline phosphatase-conjugate was from Southern Biotechnology Associates (Birmingham, Ala.). The Mt α and Mt β antibodies raised against the α and β subunits, respectively, of the *M. thermophila* 20S proteasome were described previously (156). Clones containing *M. jannaschii* open reading frames MJ0591, MJ1237, and MJ0591 were obtained from The Institute for Genomic Research (Rockville, Md.).

Strains and Media

Bacterial and archaeal strains used in this study are listed in Table 3. *E. coli* strains were grown in Luria Broth (LB), or LB supplemented with an appropriate antibiotic, including 34 mg chloramphenicol, 50 mg spectinomycin, 12.5 mg tetracycline, 100 mg ampicillin, and/or 50 mg kanamycin per liter unless otherwise indicated. The ER1647 cells used as a host strain for λ phage were grown in LB supplemented with tetracycline, 0.2% maltose, and 10 mM $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$. The BM25.8 cells used for Cre/lox-mediated subcloning of plasmids from λ BlueStar phage DNA were grown in LB supplemented with chloramphenicol, kanamycin, 0.2% maltose, and 10 mM MgSO_4 . The XL-1 Blue MRF' strain containing *M. maripaludis* shuttle vectors pWLG30 or pWLG40 was grown in low-salt LB (0.5% NaCl) containing 60 mg ampicillin per liter according to Gardner and Whitman (63). The *M. jannaschii* strain JAL-1 was grown to mid-log phase in MSH medium under a 80:20 H_2 : CO_2 gas mix at 40 lb/in² according to Boone *et al.* (20). The *M. maripaludis* strain JJ was grown in either MG medium (20), McC, or McN medium (257). Liquid cultures were maintained in 80:20 H_2 : CO_2 gas mix at 40 lb/in² for liquid cultures and at 15 lb/in² for plates. The *M. maripaludis* strains containing the pWLG30 or pWLG40 expression vectors were grown in McC media containing 2.5 mg per liter puromycin.

DNA and Transformation Techniques

The concentration of genomic and plasmid DNA was determined by measuring absorbance at 260 nm using a DU640 spectrophotometer (Beckman, Schaumburg, Ill). Plasmids from *E. coli* and *M. maripaludis* strains were isolated using the Quantum Plasmid Mini-Prep kit (Bio-Rad, Hercules, Calif.) with the following modifications for

TABLE 3. Bacterial and Archaeal strains used in this study

Strain	Description and/or Genotype	Source
<i>E. coli</i> BL21(DE3)	F' <i>ompT[lon]hsdS_B</i> ($r_B^- m_B^-$) with DE3, a λ prophage carrying the T7 RNA polymerase gene (238)	Stratagene
<i>E. coli</i> BM25.8	<i>supE thi</i> $\Delta(lac-proAB)$ [F' <i>traD36 proA⁺B⁺lac^fZ</i> Δ M15] λ imm ⁴³⁴ (kan ^R) P1 (Cm ^R) <i>hsdR</i> ($r_{k12}^- m_{k12}^+$)	Novagen
<i>E. coli</i> DH5 α	F' <i>endA1 hsdR17</i> ($r_K^- m_K^+$) <i>supE44 thi-1 recA1 gyrA</i> (Nal ^I) <i>relA1</i> $\Delta(lacIZYA-argF)U169$ <i>deoR</i> (ϕ 80 <i>dlacI</i> $\Delta(lacZ)M15$)	Life Technologies
<i>E. coli</i> ER1647	F' <i>fhuA2</i> $\Delta(lacZ)r1$ <i>supE44 recD1014 trp-31 mcrA1272::Tn10</i> (Tc ^R) <i>his-1 rpsL104</i> (str ^R) <i>xyl7 mtI 2 metB1</i> $\Delta(mcrC-mrr)102::Tn10$ (Tc ^R) <i>hsdS</i> ($r_{k12}^- m_{k12}^+$)	Novagen
<i>E. coli</i> MORPH <i>mutS</i>	$\Delta(mcrA)$ $\Delta(mcrBC-hsdSMR-mrr)$ <i>endA1 supE thi gyrA relA1 lac mutS::Tn10</i> (Tet ^r) [F' <i>proAB lac^f lacZ</i> Δ M15 Tn5]	5 Prime $\rightarrow$ 3 Prime
<i>E. coli</i> TB-1	F' <i>ara</i> $\Delta(lac-proAB)$ <i>rpsL</i> (Str ^R) [ϕ 80 <i>dlac</i> Δ (<i>lacZ</i>)M15] <i>thi hsdR</i> ($r_K^- m_K^+$)	New England Biolabs
<i>E. coli</i> XL-1 Blue MRF'	$\Delta(mcrA)183$ $\Delta(mcrBC-hsdSMR-mrr)173$ <i>endA1 supE44 thi-1 recA1 gyrA96 relA1 lac</i> [F' <i>proAB lac^f Z</i> Δ M15 Tn10 (Tet ^r)]	Stratagene
<i>M. jannaschii</i> JAL-1	Wild type (108)	David Boone, Portland State Univ.
<i>M. maripaludis</i> JJ	Wild type (106)	David Boone, Portland State Univ.

M. maripaludis. After 1 ml of a mid-log culture was centrifuged, the pellet was resuspended in McC medium rather than the resuspension buffer provided by the manufacturer to prevent premature lysis of the cells. The *E. coli* strains listed in Table 3 were transformed with plasmid DNA by either chemical transformation or electroporation. The *M. maripaludis* strain JJ was transformed using techniques according to Tumbula *et al.* (260) with modifications as per W. L. Gardner (62). For all other molecular biology procedures, standard techniques were employed (206).

Protein Techniques and Enzyme Assays

Protein concentrations were determined by the bicinchoninic acid method (232) (Pierce, Rockford, Ill.) and/or the Coomassie blue dye-binding method (22) (BioRad, Hercules, Calif.) using bovine serum albumin (BSA) as the standard. The α -amino groups generated by peptide-bond hydrolysis were measured by fluorescence at pH 6.8 using fluorescamine with methionine as a standard (261). Fluorescence was measured using an Aminco Bowman Series 2 Luminescence Spectrometer (Spectronic Instruments, Rochester, NY) with an excitation wavelength of 370 nm and an emission wavelength of 480 nm. Peptide-hydrolyzing activity was assayed by fluorimetric measurement of the release of 7-amino-4-methylcoumarin or the colorimetric measurement of the release of β -naphthylamine using synthetic peptide substrates as previously described (156, 157). Nucleotidase activity was assayed by measuring the release of inorganic phosphate (P_i) using malachite green and ammonium molybdate with modifications as previously described (132, 142). For analysis of various nucleotide-hydrolyzing activities, enzyme (2 μ g per ml) was incubated with 1 mM nucleotide in 25 mM TES at pH 8.0 with 100 mM NaCl and 10 mM $MgCl_2$ (TES-NaCl- $MgCl_2$ buffer) at 80°C for MjPAN or 65°C for

MjPAN($\Delta 1-73$). Specific activities are reported as nanomoles of product $\cdot \text{min}^{-1} \cdot \text{mg}^{-1}$ of protein. For determination of kinetic parameters, enzyme (4 μg per ml) was incubated with ATP or CTP at 0.05, 0.1, 0.25, 0.5, and 1 mM using similar assay conditions. Initial velocity was determined by measuring the release of P_i in triplicate using four time points from 0 to 6 min. Initial velocity was plotted as a function of substrate concentration by the Lineweaver-Burk method (39). Similar results were obtained using Michaelis-Menten and Eadie-Hofstee plots (39). The effects of nucleotide diphosphates were measured by preincubating the enzyme (2 μg per ml) with or without 1 mM CDP or ADP for 15 min in TES-NaCl-MgCl₂ buffer at 21°C before addition of 1 mM ATP or CTP and then assayed as described above. The pH optimum of enzymes was determined using the following buffers at 25 mM: MES (pH 5 to 7), TES (pH 7 to 8), Tricine (pH 8 to 8.8), and CAPSO (pH 8.9 to 10).

Molecular Mass Determinations

Molecular masses of the purified PAN and 20S proteasome proteins were determined by reducing and denaturing sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) using 12% polyacrylamide gels (141) which were stained with Coomassie blue R-250. The molecular weight standards for SDS-PAGE were phosphorylase b (97.4 kDa), serum albumin (66.2 kDa), ovalbumin (45 kDa), carbonic anhydrase (31 kDa), trypsin inhibitor (21.5 kDa), and lysozyme (14.4 kDa) (BioRad). Native molecular masses of the complexes were determined by applying purified proteins to a Superose 6 HR 10/30 column. For the PAN protein from *M. jannaschii*, the column was equilibrated in 20 mM Tris buffer containing 150 mM NaCl at pH 8.0. The molecular mass standards were applied to the same column equilibrated in 50 mM Tris-

HCl buffer with 150 mM NaCl and 1 mM DTT at pH 7.2. These standards included: serum albumin (66 kDa), alcohol dehydrogenase (150 kDa), β -amylase (200 kDa), apoferritin (443 kDa), and thyroglobulin (669 kDa).

Cloning the *M. jannaschii* Proteasome Genes

Plasmid pAMJHW03 (26) containing the open reading frame MJ1176 encoding the putative PAN protein was obtained from the American Type Culture Collection (ATCC) (Rockville, MD). Polymerase chain reaction (PCR) was used to amplify DNA fragments carrying the *M. jannaschii* pan ORF and contained 10 mM Tris-HCl buffer at pH 8.8 containing 10 mM KCl, 10 mM $(\text{NH}_4)_2\text{SO}_4$, 2 to 6 mM MgSO_4 , 0.1% Triton X-100, 250 μM dNTPs, 1.5 μM each primer and 2,500 ng double stranded plasmid DNA template, and 0.02 U/ μl Vent (exo⁺) DNA polymerase with a Gene Cyclor (BioRad). Oligonucleotide primers and template plasmid DNA are presented in Table 4. The DNA fragments generated by PCR amplification were cloned into expression vectors by either: 1) phosphorylating the PCR fragment with T4 polynucleotide kinase, ligating the fragment into the *Hinc*II site of pUC19, and then ligating the pUC19 cloned fragment into pET24b using *Nde*I and *Eco*RI or 2) directly ligating the PCR fragment into pET24b or pET15b using *Nde*I, *Eco*RI, or *Xho*I (Table 4). The DNA sequences of the fragments generated by PCR amplification were verified by DNA sequence analysis of both strands of plasmids pMjPAN3, pMjPAN6, pMjPAN8, pJAM735, and pJAM736.

TABLE 4. Plasmids used in this study.

Plasmid (stock #)	Phenotype or genotype; oligonucleotides for PCR amplification or site-directed mutagenesis	Source
pCR-TOPO	Km ^R ; cloning vector	Invitrogen
pET24b	Km ^R ; expression vector	Novagen
pET15b	Ap ^R ; expression vector	Novagen
pWLG40 + <i>lacZ</i>	Ap ^R ; Pur ^R ; expression shuttle vector	(63)
pSJS1240	Sp ^R ; <i>E. coli ileX</i> and <i>argU</i>	(124)
pAMJL73	Ap ^R ; 2.385-kb of <i>M. jannaschii</i> genomic DNA bases 524549 to 526933 of L77117 in pUC18; carries complete ORF MJ0591 (α subunit of 20S proteasome)	ATCC
pAMJKH57	Ap ^R ; 2.036-kb of <i>M. jannaschii</i> genomic DNA bases 1182020 to 1184055 of L77117 in pUC18; carries complete ORF MJ1237 (β subunit of 20S proteasome)	ATCC
pAMJHW03	Ap ^R ; 2.157-kb of <i>M. jannaschii</i> genomic DNA bases 1114515 to 1116671 of L77117 in pUC18; carries complete ORF MJ1176 (proteasome-activating nucleotidase)	ATCC
pMjPAN1 (#340)	Ap ^R ; <i>NdeI</i> to <i>EcoRV</i> 1.6-kb of <i>M. jannaschii</i> genomic DNA ligated into the <i>HincII</i> site of pUC19; carries 84-bp upstream and 191-bp downstream of the complete MJ1176	this study
pMjPAN2 (#447)	Km ^R ; 1.6-kb <i>EcoRI</i> -to- <i>HindIII</i> fragment of pMjPAN1 ligated into pET24b; produces PAN and PAN(Δ 1-73) proteins	this study
pMjPAN3 (#487)	Ap ^R ; 1,312-bp fragment generated by PCR amplification using pAMJHW03; carries the complete MJ1176 ligated into pUC19 using <i>HincII</i> ; 5'-ccgcccc <i>catatg</i> gtttttgaagaatt-3' and 5'-gcgcgc <i>gaattc</i> tatatctgtagagaa-3' (<i>NdeI</i> and <i>EcoRI</i> sites indicated)	this study
pMjPAN5 (#493)	Ap ^R ; 1302-bp <i>NdeI</i> -to- <i>EcoRI</i> fragment of pMjPAN3 ligated into pET15b; produces full-length PAN protein with His ₆ at N terminus and 39-kDa PAN(Δ 1-73)	this study
pMjPAN6 (#505)	Ap ^R ; site-directed mutation of pMjPAN3 which disrupts the internal Shine Dalgarno site; 5'-gagcttgaac <i>gtgaa</i> aatttacagtaagt-3'; altered bases indicated in italics	this study

TABLE 4. Cont.

Plasmid (stock #)	Phenotype or genotype; oligonucleotides for PCR amplification or site-directed mutagenesis	Source
pMjPAN7 (#511)	Km ^R ; 1302-bp <i>NdeI</i> -to- <i>EcoRI</i> fragment of pMjPAN6 which carries the complete MJ1176 with the internal translation mutation ligated into pET24b; produces PAN of 47 kDa	this study
pMjPAN8 (#528)	Km ^R ; 1083-bp <i>NdeI</i> -to- <i>EcoRI</i> fragment generated by PCR amplification using pAMJHW03; carries MJ1176 with a 219-bp 5'-deletion (N-terminal 73-residue truncation) ligated into pET24b; produces truncated PAN(Δ 1-73) only; 5'- <u>tccatatg</u> aaagaaaatgagatttaa-3' and 5'-gcggcga <u>aat</u> ttatctgtagagaa-3'; <i>NdeI</i> and <i>EcoRI</i> sites indicated	this study
pJAM721 (#617)	Ap ^R ; contains ~15-kbp insert of genomic DNA from <i>M. maripaludis</i> obtained from automatic Cre-lox mediated subcloning of λ BlueStar vector; carries entire 1220-bp <i>Mmp</i> an gene and upstream and downstream regions	this study
pJAM727 (#648)	Km ^R ; 1220-bp <i>NdeI</i> -to- <i>Bam</i> HI fragment generated by PCR amplification using pJAM721 as a template; carries entire <i>Mmp</i> an gene ligated into pET24b; produces MmPAN protein of 45-kDa; 5'-gacgcaac <u>catatg</u> agttaccagatg-3' and 5'-atagt <u>ggatc</u> ttatccgtacattac-3'; <i>NdeI</i> and <i>Bam</i> HI sites indicated	this study
pJAM728 (#650)	Ap ^R ; 1220-bp <i>NdeI</i> -to- <i>Bam</i> HI fragment generated by PCR amplification using pJAM721 as a template; carries entire <i>Mmp</i> an gene ligated into pET15b; produces 45-kDa MmPAN protein with His ₆ tag at N-terminus; 5'-gacgcaac <u>catatg</u> agttaccagatg-3' and 5'-atagt <u>ggatc</u> ttatccgtacattac-3'; <i>NdeI</i> and <i>Bam</i> HI sites indicated	this study
pJAM730 (#655)	Km ^R ; 1220-bp <i>NdeI</i> -to- <i>XhoI</i> fragment generated by PCR amplification using pJAM721 as a template; carries entire <i>Mmp</i> an gene ligated into pET24b; produces 45-kDa MmPAN protein with His ₆ tag at C-terminus; 5'-gacgcaac <u>catatg</u> agttaccagatg-3' and 5'-ctagtagt <u>ctc</u> gagtcctgacattac-3'	this study

TABLE 4. Cont.

Plasmid (stock #)	Phenotype or genotype; oligonucleotides for PCR amplification or site-directed mutagenesis	Source
pJAM735 (#747)	Km ^R ; <i>Nde</i> I-to- <i>Xho</i> I fragment generated by PCR amplification using pMjPAN1 as a template; carries entire unmodified <i>Mjpan</i> gene ligated into pET24b; produces 47-kDa MjPAN protein with His ₆ tag at C-terminus; 5'-ccgcgcatatggttttgaagaatt-3' and 5'tagtgatctcgagctctgtagagaac-3'; <i>Nde</i> I and <i>Xho</i> I sites indicated	this study
pJAM736 (#749)	Km ^R ; <i>Nde</i> I-to- <i>Xho</i> I fragment generated by PCR amplification using pMjPAN7 as a template; carries entire <i>Mjpan</i> gene with the internal translation mutation ligated into pET24b; produces 47-kDa MjPAN protein with His ₆ tag at C-terminus; same primers used as for construction of pJAM735	this study
pJAM737 (#769)	Ap ^R ;Pur ^R ; <i>Pst</i> I-to- <i>Bgl</i> II fragment generated by PCR using pJAM727 as a template; carries entire <i>Mmp</i> an gene (plus the T ₇ terminator sequence from pET24b) ligated into pWLG40; produces 45-kDa MmPAN protein in recombinant <i>M. maripaludis</i> ; 5'-ggagtactgcagatgagttaccag-3' and 5'-catgaagatctgcttaatgcgcgcg-3'; <i>Pst</i> I and <i>Bgl</i> II sites indicated	this study
pJAM738 (#754)	Ap ^R ;Pur ^R ; <i>Pst</i> I-to- <i>Bgl</i> II fragment generated by PCR using pJAM730 as a template; carries entire <i>Mmp</i> an gene (plus the T ₇ terminator sequence from pET24b) ligated into pWLG40; produces 45-kDa MmPAN protein with His ₆ tag at C-terminus in recombinant <i>M. maripaludis</i> ; same primers used as for construction of pJAM737	this study
pJAM739 (#830)	Ap ^R ;Pur ^R ; <i>Pst</i> I-to- <i>Bgl</i> II fragment generated by PCR using pJAM728 as a template; carries entire <i>Mmp</i> an gene (plus the T ₇ terminator sequence from pET24b) ligated into pWLG40; produces 45-kDa MmPAN protein with His ₆ tag at N-terminus in recombinant <i>M. maripaludis</i> ; 5'-aaactgcagatgggcagcagccatc-3' and 5'-acatgaagatctcgggcctcttgcg-3'; <i>Pst</i> I and <i>Bgl</i> II sites indicated	this study

TABLE 4. Cont.

Plasmid (stock #)	Phenotype or genotype; oligonucleotides for PCR amplification or site-directed mutagenesis	Source
pJAM740 (#771)	Ap ^R ;Pur ^R ; <i>Nsi</i> I-to- <i>Bgl</i> II fragment generated by PCR using pJAM735 as a template; carries entire unmodified <i>Mjpan</i> gene (plus the T ₇ -terminator sequence from pET24b) ligated into pWLG40; produces 47-kDa MjPAN protein with His ₆ tag at C-terminus in recombinant <i>M. maripaludis</i> ; 5'-gcggatgcattggttttgaagaatt-3' and 5'-catgaagatctgcttaatgcgccgc-3'; <i>Nsi</i> I and <i>Bgl</i> II sites indicated	this study
pJAM741 (#774)	Ap ^R ;Pur ^R ; <i>Nsi</i> I-to- <i>Bgl</i> II fragment generated by PCR using pJAM736 as a template; carries entire <i>Mjpan</i> gene with the internal translation mutation (plus the T ₇ -terminator sequence from pET24b) ligated into pWLG40; produces 47-kDa MjPAN protein with His ₆ tag at C-terminus in recombinant <i>M. maripaludis</i> ; same primers used as for pJAM740	this study
pJAM742 (#826)	Ap ^R ;Pur ^R ; <i>Nsi</i> I-to- <i>Bgl</i> II fragment generated by PCR using pMjPAN7 as a template; carries entire <i>Mjpan</i> gene with the internal translation mutation (plus the T ₇ -terminator sequence from pET24b) ligated into pWLG40; produces 47-kDa MjPAN protein in recombinant <i>M. maripaludis</i> ; same primers used as for pJAM740	this study
pJAM743 (#828)	Ap ^R ;Pur ^R ; <i>Nsi</i> I-to- <i>Bgl</i> II fragment generated by PCR using pMjPAN5 as a template; carries entire unmodified <i>Mjpan</i> gene (plus the T ₇ -terminator sequence from pET24b) ligated into pWLG40; produces 47-kDa MjPAN protein with His ₆ tag at N-terminus in recombinant <i>M. maripaludis</i> ; 5'-agataatgcattgggcagcagccatc-3' and 5'-acatgaagatctcgccctcttgcg-3'; <i>Nsi</i> I and <i>Bgl</i> II sites indicated	this study
pJAM745 (#800)	Km ^R ; 750-bp fragment generated by PCR amplification using <i>M. maripaludis</i> genomic DNA as a template; carries 333-bp upstream and 417-bp of the 5' region of <i>Mmpn</i> ligated into pCR-TOPO; 5'-cctcttttcgcagattgcattg-3' and 5'gattccattgccatagctcggtg-3'	this study

Site-directed Mutagenesis

Oligonucleotide-directed mutagenesis was performed using the MORPH site-specific plasmid DNA mutagenesis kit as recommended by the supplier (5 Prime→ 3 Prime). Double-stranded plasmid pMjPAN1 was used as a template for annealing of the oligonucleotide 5'-GAGCTTGAAC***CGTGA***AAATTTACAGTTAATG-3', where the bold italicized bases are the conservative codon replacements used to alter the internal ribosomal binding site of ORF MJ1176. Plasmid DNA was treated with *DpnI* to destroy double-stranded methylated non-mutagenized plasmids. Mutagenized plasmids were transformed into *E. coli* strain MORPH *mutS* (Table 3), and mutations were confirmed by DNA sequence analysis of plasmid pMjPAN6 (Table 4).

Expression of the *M. jannaschii* and *M. maripaludis* pan Genes in *E. coli*

The *M. jannaschii* and *M. maripaludis* pan genes which encode the PAN protein were expressed in *E. coli* BL21(DE3) using the bacteriophage T₇ RNA polymerase-promoter system with plasmids listed in Table 4 (Novagen, Madison, WI) (238). An *E. coli* host strain containing the *ileX* and *argU* genes, encoding tRNA_{AUA} and tRNA_{AGA/AGG}, on plasmid pSJS1240 (123), was transformed with appropriate expression plasmids to maximize the yield of recombinant *M. jannaschii* and *M. maripaludis* proteins. Transformed cells were inoculated into LB medium supplemented with appropriate antibiotics and grown at 37°C and 200 RPM until the cells reached mid-log phase ($A_{600nm} \sim 0.7$ units). The T₇ RNA polymerase-dependent transcription was induced with 0.4 mM IPTG for 2 h. Cells were harvested by centrifugation at 5,000 x *g* for 15 min at 4°C, and the cell material was either used immediately or stored at -70°C.

Purification of MjPAN Proteins from Recombinant *E. coli*

For purification of N-terminal poly-histidine-tagged MjPAN (His₆-MjPAN), *E. coli* (pMjPAN5) cells were thawed in 6 volumes (wt/vol) of 20 mM Tris buffer at pH 7.9 containing 5 mM imidazole and passed through a French pressure cell at 20,000 lb/in². This was followed by centrifugation at 16,000 x g for 20 min at 4°C. The extract was applied to a Ni²⁺-Sephacel column (4.8 by 0.8 cm) (Pharmacia) equilibrated with 20 mM Tris buffer at pH 7.9 containing 5 mM imidazole and then washed with 10 ml of 20 mM Tris buffer at pH 7.9 containing 60 mM imidazole. The His₆-MjPAN was eluted with 10 ml of 20 mM Tris buffer at pH 7.9 containing 500 mM imidazole, pooled, and dialyzed against 25 mM Tris-pH 7.5 buffer. Fractions with ATP hydrolyzing activity at 60°C were pooled and determined to be pure by reducing SDS-PAGE (141).

For purification of MjPAN proteins without the His₆-tag, cell-free lysate was generated using *E. coli* (pMjPAN2, pMjPAN7, or pMjPAN8) cells as described above with 20 mM Tris buffer at pH 8.0 containing 1 mM DTT (Tris-pH 8.0-buffer) and passed through a French pressure cell at 20,000 lb/in². The lysate was heated to 85°C for 15 min, chilled to 0°C for 15 min, and centrifuged at 10,000 x g for 30 min at 4°C which removed the majority of *E. coli* proteins. Samples were applied to a HiLoad Q-Sepharose 26/10 column (2.5 by 28.5 cm) (Pharmacia) equilibrated in Tris-pH 8.0-buffer, and the proteins were eluted with a linear NaCl gradient (0 to 500 mM NaCl in 240 ml of Tris-pH 8.0-buffer) at 375 mM NaCl. Fractions which hydrolyzed ATP at 65°C were pooled, concentrated by dialysis against PEG8000, and applied to a Superose 6 HR 10/30 column equilibrated with Tris-pH 8.0-buffer with 150 mM NaCl and 10 % glycerol. Protein fractions with ATPase activity were pooled and applied to a QHyperD (BioRad) or

MonoQ HR 5/5 column (Pharmacia) equilibrated with Tris-pH 8.0-buffer with 100 mM NaCl. The proteins were eluted with a linear NaCl gradient (100 to 500 mM NaCl in 10 ml of Tris-pH 8.0-buffer) at 370 mM NaCl. Fractions which hydrolyzed ATP at 65EC were pooled. Full-length MjPAN protein was further purified by application to a CHT5-IB hydroxyapatite column (BioRad) equilibrated with 10 mM sodium phosphate buffer at pH 7.5 containing 1 mM DTT and 10% glycerol (sodium phosphate-pH 7.5 buffer) and elution with a linear gradient (10 to 250 mM sodium phosphate-pH 7.5 buffer in 30 ml). The MjPAN protein, which eluted at 225 mM sodium phosphate-pH 7.5 buffer, was dialyzed against 50 mM Tris buffer at pH 7.5 containing 100 mM NaCl, 1 mM DTT, and 10% glycerol at 4°C for 18 h. The protein was then applied to a Sephadex G-25 (Sigma) column equilibrated with 50 mM Tris buffer at pH 7.5 containing 100 mM NaCl, 1 mM DTT, and 10% glycerol to remove all detectable P_i . The MjPAN and MjPAN(Δ 1-73) proteins were determined to be pure by reducing SDS-PAGE (141) and stored under liquid N_2 without apparent loss of activity.

Purification of Genomic DNA from *M. maripaludis*

A 100 ml late log culture of *M. maripaludis* in MG medium was centrifuged at 5,000 x g for 10 minutes at 4° C. The cell pellet was then resuspended in 5 ml TE/salts buffer (10 mM Tris, 100 mM EDTA, 5 mM $MgCl_2$, and 50 mM NaCl, pH 8.0) containing the same concentration of salt as in the growth media. Genomic DNA was then isolated following procedures according to K. R. Sowers (235).

Southern Analysis of *M. maripaludis* Genomic DNA

Genomic DNA from *M. maripaludis* was digested with several restriction enzymes, including *EcoRV*, *XbaI*, and *SmaI*, separated by 0.8% agarose gel

electrophoresis, and transferred for 16 h by upward capillary action to a positively-charged nylon membrane (234). After UV-crosslinking, the membranes were equilibrated in hybridization buffer (5X SSC, 0.1% N-lauroylsarcosine, 0.02% SDS, 1% blocking reagent) for 4h at 68°C, where 20X SSC solution consists of 3M NaCl and 0.3M sodium citrate at pH 7.0. The blot was then probed with the 1071-bp *M. jannaschii pan* ($\Delta 1-73$) gene fragment (Mjpan probe), which had been random-prime labeled with DIG-dUTP following procedures recommended by the manufacturer (Roche Biochemicals). Hybridization was performed with the Mjpan probe at a final concentration of 20 ng per ml for 14 h at 68° C. The membrane was then washed in 2X SSC buffer containing 0.1% SDS (2 x 5 min), and in 0.5X SSC buffer containing 0.1% SDS (2 x 15 min) at 68° C. Signals were then visualized colorimetrically using alkaline phosphatase-conjugate antibody raised against DIG according to instructions provided by the supplier (Roche Biochemicals).

Construction of a *M. maripaludis* Random Genomic λ Library

A random genomic λ library was constructed using *M. maripaludis* genomic DNA which had been partially digested with the endonuclease *Sau3AI*. The procedures were as recommended by the manufacturer (Novagen) with the following modifications. After partial fill-in of the *Sau3AI*-digested cohesive ends, the DNA was loaded onto a 0.8% agarose gel and size fractionated to isolate 7- to 20-kbp fragments for packaging into λ BlueStar phage. After excision of the agarose gel piece containing fragments of the desired size, the DNA was electroeluted according to Peloquin and Platzer (187). The DNA fragments were then ligated into λ BlueStar *XhoI*-digested vector arms, added to λ phage packaging extract, and plated onto the host strain ER1647 (Table 3). After

amplification of the library according to the procedures of the manufacturer (Novagen), the λ library was stored in SM buffer (50 mM Tris at pH 7.5 containing 100 mM NaCl, 8 mM MgSO₄, and 0.01% gelatin) in 0.5 ml aliquots containing 7% DMSO at -70° C.

Isolation and Cloning of the *Mmp* Gene from *M. maripaludis*

After plating the phage from the *M. maripaludis* genomic λ library, the phage were grown for 14 h at 37° C on LB and then plates were chilled at 4° C for 1 h. Plaques were then lifted onto nylon membranes, treated with denaturation solution (1.5M NaCl, 0.5M NaOH) for 5 min, and neutralized with 1.5M NaCl and 1M Tris (pH 7.5) for 15 min. After treatment with 2X SSC for 15 min, the DNA was fixed to the wet membranes by UV-crosslinking. The membranes were then probed with the *Mjpan* probe, using the hybridization conditions and detection methods described above. Plasmid subclones of positive plaques were obtained by automatic Cre-lox-mediated subcloning using the host strain BM25.8 (Table 3). The BM25.8 strain expresses the Cre recombinase which will excise the λ vector at two loxP sites by site-specific recombination. This yields a pUC-derived plasmid containing the genomic insert of interest, a gene conferring ampicillin/carbenicillin resistance, and an *E. coli* origin of replication. Following the procedures recommended by the manufacturer (Novagen), the λ phage was adsorbed to host strain BM25.8 and then the host/phage mixture was plated onto LB agar plates containing 50 mg per liter carbenicillin. Plasmid replication and copy number in strain BM25.8 is very low; therefore, selected plasmid subclones were isolated and re-transformed by electroporation into the *E. coli* host strain DH5 α . After isolation of the plasmid DNA from these clones, Southern analysis was performed using identical conditions as described above for analysis of *M. maripaludis* genomic DNA.

Cloning of the Mmp_{an} Gene into *E. coli* Expression Vectors

Plasmid pJAM721 was used as a template in polymerase chain reactions to amplify DNA fragments carrying the *M. maripaludis* ORF. The PCR buffer conditions were as described for Mj_{pan} except that 55 ng of pJAM721 was used as a template. Cycling conditions were 1 min at 94°, 2 min at 50°C, and 2 min at 72° C (for 40 cycles). The DNA fragments generated by PCR amplification were cloned into the expression vectors pET15b (for N-terminal His₆-tagged construct) or pET24b (for native and C-terminal His₆-tagged constructs) using *Nde*I and *Bam*HI or *Xho*I (Table 4). The DNA sequences of the fragments generated by PCR amplification were verified by DNA sequence analysis of both strands of plasmids pJAM727, pJAM728, and pJAM730.

To verify the start site of the Mmp_{an} gene, the 5' region was amplified directly from genomic DNA to construct plasmid pJAM745. The PCR buffer conditions were the same as those described above except that 300 ng of genomic DNA was used as a template, and the annealing temperature was 55°C. After PCR, the fragments were directly cloned, via a topoisomerase I-mediated direct ligation, into the pCR-TOPO vector using procedures as recommended by the supplier (Invitrogen). The start site of the Mmp_{an} ORF was then verified by DNA sequence analysis of both strands of pJAM745 (Table 4).

Purification of MmPAN Proteins from Recombinant *E. coli*

For purification of N- and C-terminal poly-histidine-tagged MmPAN (N- and C-His₆-MmPAN) proteins, *E. coli* (pJAM728 and pJAM730) cells were thawed in 8 and 6.5 volumes (wt/vol), respectively, of 20 mM Tris buffer at pH 7.5 containing 100 mM NaCl and 5 mM imidazole and passed through a French pressure cell twice at 20,000 lb/in².

This was followed by centrifugation at $15,000 \times g$ for 20 min at 4°C . All samples and buffers were filtered through a $0.45\mu\text{m}$ filter. Samples (48.4 mg total for N-His₆-MmPAN and 74 mg total for C-His₆-MmPAN) were applied to Ni²⁺-Sepharose columns (4.8 by 0.8 cm) (Pharmacia) which had been equilibrated with 20 mM Tris buffer at pH 7.5 containing 100 mM NaCl and 5 mM imidazole and then washed with 10 ml of 20 mM Tris buffer at pH 7.5 containing 100 mM NaCl and 60 mM imidazole. The His₆-MmPAN proteins were eluted with 10 ml of 20 mM Tris buffer at pH 7.5 containing 100 mM NaCl and 500 mM imidazole and pooled. For the C-His₆-MmPAN construct only, 0.63 mg of the protein was then applied to a Superose 6 size exclusion column equilibrated with 25 mM Tris buffer at pH 7.5 containing 150 mM NaCl, 1 mM DTT, and 10% glycerol. For purification of MmPAN proteins without the His₆-tag, *E. coli* (pJAM727) cells were thawed in 7 volumes (wt/vol) of 20 mM Tris buffer at pH 7.5 containing 150 mM NaCl, 1 mM DTT, and 10% glycerol (resuspension buffer) and passed through a French pressure cell twice at $20,000 \text{ lb/in}^2$. This was followed by centrifugation at $16,000 \times g$ for 20 min at 4°C . Cell-free lysate (5 mg) was then applied to a Superose 6 HR 10/30 column which had been equilibrated in resuspension buffer. Fractions (300- to 600-kDa) which hydrolyzed ATP at 37°C were then pooled and applied to a MonoQ HR 5/5 column (Pharmacia) which had been equilibrated in 25 mM Tris buffer at pH 7.5 containing 150 mM NaCl, 1 mM DTT, and 10% glycerol. The proteins were eluted with a linear NaCl gradient from 150 mM to 1 M NaCl in 20 ml of buffer. In all purification steps, the purity of protein was measured by reducing SDS-PAGE.

Cloning of the *Mmpan* and *Mjpan* Genes into pWLG40 Expression Vectors

Plasmids pJAM727, pJAM728, and pJAM730 were used as templates for PCR amplification of the *Mmpan* ORF as well as the downstream T₇-terminator sequences contained on the pET15b and pET24b vectors. The PCR-generated fragments were directly cloned, via a topoisomerase I-mediated ligation, into pCR-TOPO vectors using procedures recommended by the supplier (Invitrogen). After DNA sequence analysis, *Pst*I-to-*Bgl*II fragments from the pCR-TOPO vectors were cloned into *Nsi*I-*Bgl*II cut pWLG40 vectors to construct the plasmids pJAM737, pJAM738, and pJAM739 (Table 4). Since the *Mmpan* gene contains an *Nsi*I site, the compatible restriction enzyme *Pst*I was used rather than *Nsi*I for cloning of the *Mmpan* gene into pWLG40. The start codon was thus 8 bp, rather than 5 bp, away from the Shine-Dalgarno sequence in the cloning region of the pWLG40 vector.

Identical procedures to those described above were used to clone the *Mjpan* ORF into the pWLG40 vector, with the following differences. The pMjPAN5, pMjPAN7, pJAM735, and pJAM736 plasmids were used as templates in PCR reactions. The *Nsi*I-to-*Bgl*II fragments were then cloned into the pWLG40 vector to generate the expression plasmids pJAM740, pJAM741, pJAM742, and pJAM743 (Table 4).

Synthesis of MmPAN and MjPAN Proteins in Recombinant *M. maripaludis* and Purification of C-terminal His₆-tagged MmPAN

M. maripaludis strains containing pJAM737, pJAM738, pJAM739, pJAM740, pJAM741, pJAM742, and pJAM743 expression plasmids (Table 4) were grown in McC medium containing 2.5 mg per liter puromycin. Expression of MmPAN and MjPAN proteins was monitored by SDS-PAGE and immunoblotting using anti-MjPAN antibodies. The purification of C-terminal His₆-tagged MmPAN was performed at 20°C

to 25°C in an anaerobic chamber (Coy Laboratory Products, Ann Arbor, Mich.) containing N₂:H₂ (19:1). All solutions were filtered through a 0.45µm filter and made anoxic as previously described (249). The *M. maripaludis* cells harboring pJAM738 (Table 4) and pWLG40 + *lacZ* (as a control) were centrifuged anaerobically at 2,500 x g for 15 min and washed two times in 20 mM Tris buffer at pH 7.4 containing 375 mM NaCl, 5 mM imidazole, and 10% glycerol (5 mM imidazole buffer). The cell suspension was then passed through a French pressure cell at 20,000 lb/in² and centrifuged at 15,000 x g for 15 min at 4°C. Cell lysate from *M. maripaludis* (pJAM738) and *M. maripaludis* (pWLG40+ *lacZ*) was then applied to separate Cu²⁺-Sephacrose columns (4.8 by 0.8 cm) (Pharmacia) equilibrated with 5 mM imidazole buffer. The columns were washed with 10 ml of 20 mM Tris buffer at pH 7.4 containing 375 mM NaCl, 60 mM imidazole, and 10% glycerol. The His₆-MmpAN proteins were eluted with 10 ml of 20 mM Tris buffer at pH 7.4 containing 375 mM NaCl, 500 mM imidazole, and 10% glycerol. The C-His₆-MmpAN (0.45 mg) was then further purified by application to a Superose 6 size exclusion column equilibrated with 20 mM Tris buffer at pH 7.4 containing 250 mM NaCl, 5 mM DTT, and 10% glycerol. In all purification steps, the purity of protein was measured by reducing SDS-PAGE.

Partial Purification of MjPAN from *M. jannaschii*

The *M. jannaschii* strain JAL-1 was grown to mid-log phase in MSH medium under an 80:20 H₂:CO₂ gas mix at 40 lb/in² according to Boone *et al.* (20). *M. jannaschii* cultures were chilled to 0°C, and the cells were harvested aerobically by centrifugation at 5,000 x g for 15 min at 4°C. Cell material (0.5 g) was stored at -70°C. Cells were thawed in 17 volumes (wt/vol) of 50 mM Tris buffer at pH 8.0 with 150 mM NaCl, 10 mM

MgCl₂, 1 mM ATP, 10 mM β -mercaptoethanol, 10 % glycerol, and 0.1 % Triton X-100. The cell suspension was passed through a French pressure cell twice at 10,000 lb/in² and then centrifuged at 15,000 x g for 30 min at 4°C. Cell lysate (2 mg) was applied to a Superose 6 HR 10/30 column equilibrated with Tris-pH 8.0-buffer with 150 mM NaCl and 10 % glycerol. Fractions with NEM-inhibitable ATPase activity were applied to a MonoQ HR 5/5 column equilibrated with Tris-pH 8.0-buffer with 100 mM NaCl and 10% glycerol. The proteins were eluted with a linear NaCl gradient (0.1 to 1 M NaCl in 15 ml of Tris-pH 8.0-buffer). In some experiments, ion-exchange chromatography preceded gel filtration. Protein fractions were analyzed for NEM-inhibitable hydrolysis of ATP and immunologically with MjPAN antibody after denaturing and reducing 12% SDS-PAGE (90). Protein was also analyzed by immunoblot of proteins separated by non-denaturing gel electrophoresis which were transferred by capillary action to PVDF membranes using 10 mM MES buffer at pH 6.0 with 20% methanol. Western analysis was performed as described above using anti-MjPAN antibodies.

Purification of a 20S Proteasome and 2 NTPases from *M. maripaludis*:

The *M. maripaludis* cell material provided by Brian Waters and William Whitman of the University of Georgia was prepared as follows. Mid- to late-log phase *M. maripaludis* cultures ($A_{600}=0.85-0.9$) grown in McN media were harvested anaerobically using an Alfa Laval Sharples (Model T-1P) centrifuge. The cell paste was then stored anaerobically at -70° C in 60 ml Wheaton bottles which had been pre-flushed with N₂. For purification of the 20S proteasome and NTPase proteins, cells were thawed in 8 volumes (wt/vol) of 25 mM Tris buffer at pH 7.5 with 5 mM DTT and 10% glycerol. The cell resuspension was passed through a French pressure cell twice at 20,000 lb/in² and

then centrifuged at $15,000 \times g$ for 20 min at 4°C . Cell lysate (20 mg protein per ml) was applied to a Q Sepharose 26/10 column (2.5 by 28.5 cm) (Pharmacia) equilibrated in 25 mM Tris buffer at pH 7.5 with 5 mM DTT and 10% glycerol, and the proteins were eluted with a linear NaCl gradient (0 to 1 M NaCl in 400 ml buffer).

For purification of the 20S proteasome, fractions which hydrolyzed LLVY-AMC were pooled, concentrated by dialysis against PEG8000 to a concentration of 6.6 mg per ml, and applied to a Superose 6 HR 10/30 column equilibrated with 25 mM Tris buffer at pH 7.5 with 150 mM NaCl, 5 mM DTT, and 10% glycerol. Fractions with activity were pooled, dialyzed in sodium phosphate buffer at pH 7.5 containing 5 mM DTT and 10% glycerol (dialysis buffer), and applied to a CHT5-IB hydroxyapatite column (BioRad) equilibrated with the dialysis buffer. The 20S proteasome was eluted (at 500 mM sodium phosphate) with a linear gradient from 15 mM to 1M sodium phosphate-buffer containing 5 mM DTT in 40 ml. Fractions with activity were pooled and dialyzed in 25 mM Tris buffer at pH 7.5 containing 150 mM NaCl, 5 mM DTT, and 10% glycerol, and applied to a MonoQ HR 5/5 column equilibrated in the same buffer. The 20S proteasome was concentrated and eluted in 25 mM Tris buffer containing 700 mM NaCl, 5 mM DTT, and 10% glycerol.

For partial purification of the two NTPase proteins, Q Sepharose 26/10 chromatography, PEG-mediated concentration, and Superose 6 gel filtration were as described for purification of the 20S proteasome. Two fractions which hydrolyzed ATP (NTPase #1 and NTPase #2) were identified. These fractions were loaded onto a MonoQ 5/5 column which was equilibrated in 25 mM Tris at pH 7.5 containing 150 mM NaCl, 5 mM DTT, and 10% glycerol. Proteins were eluted with a linear gradient from 150 mM to

1 M NaCl in 20 ml. The NTPase #1 eluted at 275 mM NaCl, and NTPase #2 eluted at 360 mM NaCl.

Production of Antibodies against MjPAN(Δ 1-73)

The recombinant MjPAN(Δ 1-73) protein of *M. jannaschii* was applied to a reducing SDS-12% polyacrylamide gel, and a gel fragment containing the protein (300 μ g) was used to generate polyclonal antibodies in rabbits with Freund's adjuvant according to supplier (Cocalico Biologicals, Reamstown, PA).

Non-denaturing Gel Electrophoresis of PAN Proteins

The PAN complexes were separated electrophoretically under non-denaturing conditions using a 5% resolving and 4% stacking gel of 375 mM Tris-HCl at pH 8.8 (BioRad). The non-denaturing running buffer contained 25 mM Tris, 192 mM glycine, 10 mM $MgCl_2$, and 1 mM ATP at pH 8.3. The Mg^{2+} -dependent ATPase activity of proteins was visualized in the gel as previously described (189) with the following modifications. After separation of the proteins by electrophoresis, the gels were incubated with or without 10 mM NEM at 37° or 55°C in 50 mM Tris buffer at pH 8.0 containing 1 mM ATP, 10 mM $MgCl_2$, 100 mM NaCl, and 0.1 mM lead nitrate until a white lead phosphate precipitate formed. The gel was washed in deionized water (3 x 5 min) and then briefly incubated in 1% (w/v) sodium sulfide, which rapidly converted the white precipitate to a brown precipitate of lead sulfide.

Electron Microscopy of Archaeal Holo-Proteasomes

The *M. jannaschii* PAN, 20S proteasome, and equimolar mixtures of the two protein complexes purified from recombinant *E. coli* were incubated at 21°C or 50°C for 15 min in 25 mM TES buffer at pH 8.0 containing 100 mM NaCl, 10 mM $MgCl_2$, and 1

mM ATP. Protein samples were placed on 200-mesh grids coated with Formvar or carbon film, fixed with 2% cacodylate-buffered glutaraldehyde, briefly stained with 1% aqueous uranyl acetate, and treated with 0.01% bacitracin to improve spreading (81). Samples were viewed and photographed using a Zeiss EM-10CA transmission electron microscope operated at 80 kV.

Protein Sequence Determinations

The N-terminal sequences of the MjPAN and MjPAN(Δ 1-73) proteins from recombinant *E. coli* were determined from purified protein which had been separated by SDS-PAGE and electroblotted onto a polyvinylidene difluoride (Immobilon-P) membrane (Millipore). All amino acid sequencing was by automated Edman degradation (51) and was performed at the University of Florida Interdisciplinary Center for Biotechnology Research protein chemistry core facility.

DNA Sequence Determinations

The Sanger dideoxy chain-termination method was used to sequence all DNA (207) and was performed at the University of Florida Department of Microbiology and Cell Science DNA Sequencing Facility.

Protein Sequence Analysis

GenBank, EMBL, and SwissProt data bases were searched at the National Center for Biotechnology Information (NCBI; Bethesda, MD) using the BLAST network server (7). The CLUSTALW program version 1.8 (250) or Multalin was used for alignment of protein sequences accessible through NCBI. Secondary structure analysis was performed using the DSC structure prediction program (125).

RESULTS AND DISCUSSION

Synthesis of *M. jannaschii* 20S Proteasome and PAN in Recombinant *E. coli*

Three ORFs which encode putative components of an energy-dependent proteasome were identified from the genome sequence of the methanoarchaeon *M. jannaschii* (26). These ORFs were MJ0591 (encoding a putative α subunit of the 20S proteasome with 61% identity and 79% similarity to *M. thermophila* PsmA), MJ1237 (encoding a putative β subunit of the 20S proteasome with 48% identity and 71% similarity to *M. thermophila* PsmB) and MJ1176 (encoding a putative AAA subunit with 49% identity and 69% similarity to Rpt2p from *H. sapiens*). These ORFs were amplified using PCR with appropriate primers and placed into expression vectors for production of the proteins in recombinant *E. coli*. When the *E. coli* host strain BL21(DE3) was used for production of these *M. jannaschii* proteins, the yield of recombinant protein was less than 0.5% of the total cell protein as determined by SDS-PAGE and Western blot analysis. Examination of the codon composition of the three ORFs revealed that up to 12% of the codons required either tRNA_{AUA} or tRNA_{AGA/AGG} for translation, which are rare in *E. coli* (162). Therefore, a compatible plasmid pSJS1240 (Table 4) with the *E. coli ileX* and *argU* genes encoding these rare tRNAs was introduced into the host (124). In the presence of the plasmid pSJS1240, the overall yield of recombinant *M. jannaschii* protein was as high as ~10% of total soluble *E. coli* proteins indicating that the rare tRNAs limit the high level production of these proteins in *E. coli*.

M. jannaschii 20S proteasome. The genes encoding the α and β subunits of the *M. jannaschii* 20S proteasome were expressed independently in recombinant *E. coli*. The α and β proteins assembled into a 700-kDa complex based on gel filtration with an overall structure similar to other 20S proteasomes based on electron microscopy (287). The observed sizes of the α and β subunits were larger than the calculated molecular masses deduced from the *M. jannaschii* genome sequence. However, immunoblot analysis of *M. jannaschii* cell lysate confirmed that both the α and β subunits of the 20S proteasome synthesized in *M. jannaschii* were of the same apparent molecular weight as the purified recombinant proteins (Fig. 2). Therefore, MJ0591 and MJ1237 were designated the *psmA* and *psmB* genes (proteasome A and B) which is consistent with the nomenclature of the related genes from *M. thermophila* (157).

M. jannaschii proteasome-activating nucleotidase (MjPAN) protein. Expression of ORF MJ1176 in recombinant *E. coli* (pMjPAN2) resulted in the high-level production of 47-, 39-, and 30-kDa proteins which were not apparent in the absence of the plasmid (Fig. 2). These results indicated that production of MjPAN, which has a calculated molecular mass of 47-kDa, may have induced the overproduction of *E. coli* proteins such as chaperones or other stress-response proteins. Alternatively, MjPAN may be hydrolyzed or the ORF may encode proteins with different sizes. The 47- and 39-kDa proteins were stable after heating the cell lysate to 85° C. In addition, they co-purified during anion exchange, hydroxyapatite, and gel filtration chromatography as a ~600-kDa fraction which catalyzed the hydrolysis of ATP at a rate of 1,800 nmol P_i released $\cdot$ min⁻¹ $\cdot$ mg protein⁻¹ at 65° C (Fig. 3, Lane 1). These two proteins were further analyzed using N-terminal Edman degradation. The first 10 residues of the 47-kDa and 39-kDa proteins

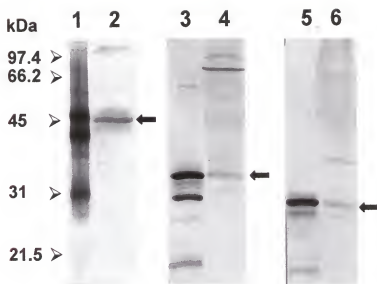


FIG. 2. Western blot analysis of *M.jannaschii* cell lysate showing that α and β subunits of the 20S proteasome and the full-length PAN protein are produced in *M. jannaschii*. Proteins were separated by 12% reducing SDS-PAGE and analyzed by Western blot. Lane 1: Cell lysate (1-2 μ g) of recombinant *E. coli* cells producing the MjPAN proteins of 47-, 39-, and 30-kDa. Lanes 2, 4, and 6: 20 μ g *M. jannaschii* cell lysate. Lanes 3 and 5: 2 μ g purified recombinant *M. jannaschii* 20S proteasome. Lanes 1 and 2 were probed with anti-MjPAN(Δ 1-73) antibody. Lanes 3 and 4 were probed with anti-Mt α . Lanes 5 and 6 were probed with anti-Mt β antibodies. Arrows indicate MjPAN, Mj α , and Mj β in lanes 2, 4, and 6 respectively. Molecular weight standards are indicated on the left.

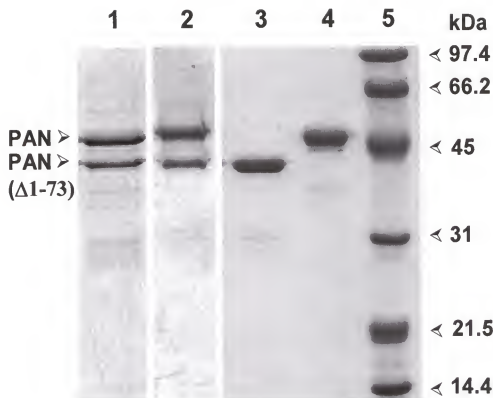


FIG. 3. *M. jannaschii* PAN proteins purified from recombinant *E. coli*. Protein (2 μ g) was separated by 12% reducing SDS-PAGE and stained with Coomassie blue. Lane 1: Full-length MjPAN and MjPAN(Δ 1-73) from recombinant *E. coli*/pMjPAN2. Lane 2: His₆-tagged MjPAN and MjPAN(Δ 1-73) from recombinant *E. coli*/pMjPAN5. Lane 3: MjPAN(Δ 1-73) from recombinant *E. coli*/pMjPAN8. Lane 4: Full-length MjPAN from recombinant *E. coli*/pMjPAN7. Lane 5: Molecular weight standards with masses indicated on right.

were found to be identical to residues 2 to 11 and 74 to 83, respectively, of the MjPAN protein sequence deduced from the MJ1176 ORF (Fig. 4). These results indicate that there is an internal translation start site recognized by *E. coli* within the full-length mRNA which accounts for the production of the shorter 39-kDa MjPAN(Δ 1-73) protein in *E. coli*. In addition, it is likely that an aminopeptidase of *E. coli* cleaved the N-terminal methionine residue of the full-length MjPAN protein; whereas the N-terminal methionine of the 39-kDa MjPAN protein was not removed by this host peptidase. Similar complex formation was observed when the MjPAN ORF was positioned in-frame to produce the full-length MjPAN protein with an N-terminal poly-histidine (His₆) tag (Table 4, pMJPAN5). Not only was a similarly truncated MjPAN protein produced in this recombinant expression system, it also co-purified with the full-length His₆-tagged MjPAN after Ni²⁺-Sepharose chromatography, even when the cell lysate was not previously exposed to 85°C (Fig. 3, Lane 2). The ATPase activity of the His₆-tagged MjPAN complex (640 nmol P_i released · min⁻¹ · mg protein⁻¹ at 65°C) was over 60% less than that of the untagged MjPAN complex described above. In addition, the tagged complex was less stable after heating to 85°C or long-term storage at 4°C compared to the untagged complex. This suggests that modification of the N-terminus of MjPAN by addition of a His₆-tag may reduce the rate of MjPAN-mediated ATP hydrolysis as well as the stability of the complex. These results also demonstrate that the MjPAN(Δ 1-73) protein which lacks 8 kDa of the N-terminus is still capable of forming a complex with full-length MjPAN even before heating of the cell lysate to 85°C.

Production and purification of a complex containing only the full-length 47-kDa PAN protein. To purify a complex containing only the full-length MjPAN protein, the

CTTTAAGAAGGAGATATACATatgggttttgaagaatttatttcaactgaattgaagaaa
 M V F E E F I S T E L K K 13
 gaaaagaaagcatttactgaagaattttaaagaagaaaaggaaataaacgataatttctaac
 E K K A F T E E F K E E K E I N D N S N 33
 ttaaaaaatgatttacttaagaggaactccaagaaaaggcaagaattgcagaattagaa
 L K N D L L K E E L Q E K A R I A E L E 53
 Oligo: gagcttgaacggtaaatattacagtta
 agtagaatcctaaaattagaattagagaaaaagagcttgaaagagagaatttacagtta
 S R I L K L E L E K K E L E R E N L Q L 73
 atg
 atgaaagaaaatgagattttaagaagagaattagatagaatgagagtcctccattgata...
M K E N E I L R R E L D R M R V P P L I...93

FIG. 4. Site-directed mutagenesis of the ribosomal binding site located within the *Mjpan* gene recognized in *E. coli*. Capitalized nucleotides are from plasmid pET24b, and lowercase nucleotides are from the *pan* gene. The Shine-Dalgarno site is double underlined. The nucleotides targeted and the oligonucleotide used for site-directed mutagenesis are in reverse type and indicated above the DNA coding region, respectively. Amino-acid residues identical to the N-terminal protein sequence determined by Edman degradation of purified 47-kDa MjPAN species are underlined, while those of the purified 39-kDa MjPAN ($\Delta 1-73$) species are underlined and italicized. Amino acid positions of the MjPAN protein deduced from the sequence of ORF MJ1176 are indicated on the right.

putative internal ribosomal binding site of the ORF (AGAGAG; Fig.4) which was recognized by *E. coli* was altered by site-directed mutagenesis at nucleotides A202C, A204T, and G207A (Fig. 4). This removed the Shine-Dalgarno consensus sequence while conserving the amino acid sequence of the MjPAN protein. After the modified gene was cloned into a pET24b vector and expressed in *E. coli* (Table 4, pMjPAN7), the 39-kDa product was not evident on either SDS-PAGE or Western blots. A 550-kDa complex containing only the full-length 47-kDa PAN protein was then purified to apparent homogeneity (Fig.3, Lane 4).

Production and purification of a complex containing only the 39-kDa truncated MjPAN(Δ 1-73) protein. In an attempt to separately produce the N-terminally truncated 39kDa protein, the ribosomal binding site upstream from the MJ1176 coding sequence on the plasmid pMjPAN7 (Table 4) was removed leaving the full-length gene intact. However, based on a Western blot, a small amount of the full-length 47-kDa protein was still produced. Therefore, the 219-bp of DNA coding for the 8-kDa N-terminus of MjPAN was removed by PCR amplification leaving only the region of the gene that produces the 39-kDa species. This modified gene was then placed into pET24b with the Met⁷⁴ ATG start codon positioned 8 bp downstream from the ribosomal binding site on the expression plasmid pMjPAN8 (Table 4; Fig. 4). The truncated MjPAN(Δ 1-73) also purified as a large molecular weight complex (~500kDa) based on gel filtration and non-denaturing native-gel analysis (Fig.5), indicating that the N-terminus is not required for oligomerization of the complex. Thus, the full-length MjPAN and MjPAN(Δ 1-73) proteins were produced separately and purified to homogeneity, as determined by reducing SDS-PAGE (Fig. 3, Lanes 3 and 4).



FIG. 5. Non-denaturing electrophoretic separation of PAN proteins. Lane 1: apoferritin (443 kDa). Lane 2: thyroglobulin (669 kDa). Lane 3: Truncated PAN(Δ 1-73) of 500-kDa based on gel filtration purified from recombinant *E. coli*. Lanes 4, 5, 7, and 8: Full-length PAN (550-kDa based on gel filtration) purified from recombinant *E. coli*. Lane 6: Native PAN partially purified from *M. jamnaschii*. Lanes 1-4 were stained with Coomassie Blue. Lanes 5 and 6 were analyzed immunologically with anti-PAN (Δ 1-73) antibodies. Lanes 7 and 8 were tested for the generation of P_i from ATP in the absence (lane 7) and presence of 10 mM NEM (lane 8).

Analysis of *M. jannaschii* Cell-Free Extract and Partial Purification of Native MjPAN

Cell lysate of *M. jannaschii* grown on H₂ and CO₂ at 65° C was analyzed by immunoblot. Western blots using anti-Mt α and anti-Mt β raised against the subunits of the *Methanosarcina thermophila* 20S proteasome demonstrated that both the α and β subunits of the 20S proteasome are synthesized in *M. jannaschii* (Fig. 2, Lanes 3-6). In addition, immunoblot analysis using anti-MjPAN (Δ 1-73) antibody demonstrated that ORF1176 encoding the *pan* gene is synthesized in *M. jannaschii* as only the full-length MjPAN protein of 47 kDa when grown under these conditions (Fig. 2, Lanes 1,2). Partial purification of PAN from *M. jannaschii* cell lysate also revealed that this protein is associated in a distinct complex (Fig. 5, Lane 6) of 550-kDa as determined by gel filtration. Similarly, the MjPAN and MjPAN (Δ 1-73) proteins purified from recombinant *E. coli* each form large molecular-weight complexes (Fig. 5, Lanes 3-5) of 550 and 500 kDa, respectively, based on gel filtration. It remains to be determined if the MjPAN purified from *M. jannaschii* is a homooligomer or associates with additional proteins.

Structure of Recombinant MjPAN and MjPAN(Δ 1-73)

Based on gel filtration, both MjPAN and MjPAN(Δ 1-73) proteins synthesized in recombinant *E. coli* formed homooligomeric complexes of 550 and 500 kDa, respectively. These results indicate that the MjPAN protein forms a complex of approximately 12 subunits and that the first 73 amino acids of the N-terminus are not required for this association.

Both the full-length MjPAN and truncated MjPAN(Δ 1-73) protein complexes were stable during purification in the absence of ATP but were readily inactivated when incubated at 65° C in glass tubes. The related ClpX ATPase protein of *E. coli* also has a

strong tendency to bind to glass (288). The high stability in the absence of ATP is in contrast to other related AAA proteins, including ClpA, ClpX, ClpQ (HslU), and NEM-sensitive fusion protein (NSF), which have an absolute requirement of ATP for complex formation and stability (82, 89, 118, 159). Full-length MjPAN migrated on non-denaturing polyacrylamide gels as a single-distinct band even after preincubation at 65° C (Fig. 5, Lane 4). This distinct band catalyzed the hydrolysis of ATP and was inhibited by NEM, as determined by in-gel staining for the generation of inorganic phosphate (Fig. 5, Lanes 7 and 8). In contrast to the related *E. coli* HslU AAA-ATPase (294), preincubation of MjPAN or MjPAN(Δ 1-73) with NEM at 65° C prior to electrophoresis did not result in dissociation of the complex. Although the truncated MjPAN(Δ 1-73) protein also migrated as a single band (Fig. 5, Lane 3), the majority of MjPAN(Δ 1-73) protein did not enter the non-denaturing gel, suggesting that it is susceptible to aggregation under these electrophoretic conditions. However, the complex is stable when maintained in optimal buffer and salt conditions based on gel filtration.

Although the number of MjPAN subunits per complex remains to be established, the gel filtration results suggest that the complex may be dodecameric. The conservation of the NSF-D2 oligomerization domain and the potential 12 subunit configuration initially suggested that the MjPAN protein may form double-stacked hexameric rings similar to other AAA proteins including NSF (89), VAT from *T. acidophilum* (181), and ARC from *R. erythropolis* (289). Transmission electron micrographs of the recombinant protein, however, showed that MjPAN forms particles of approximately 14 nm in diameter which are somewhat ring-like but are not distinct symmetrical hexameric complexes (Fig. 6). These results suggest that MjPAN may not be stable during

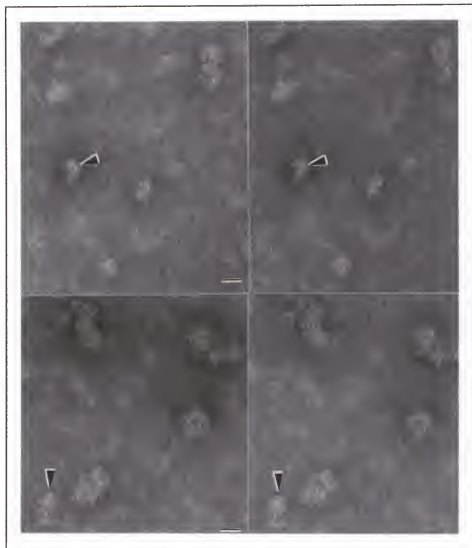


FIG. 6. Transmission electron micrograph stereo pairs of *Mjannaschii* PAN purified from recombinant *E. coli*. Several views of the comma-shaped MjPAN assembly are seen. Arrowhead on top panel indicates side view in which MjPAN most resembles a comma. Arrowhead in bottom panel points to a more face-on view in which the head of the comma is at top. Distinct individual subunits are seen as white globular components of each assembly. Some larger clusters most likely represent random dimeric and trimeric associations created when the sample was dried for TEM observation. Bars=10 nm.

preparation for microscopy using uranyl acetate as a negative stain or does not form a symmetrical complex which is easily visualized using these methods.

Biochemical Properties of the MjPAN Complexes

Past studies on AAA family members, including the yeast Yhs4p/Rpt2p (153) and *E. coli* ClpA (216, 217) and ClpB (290) proteins suggested that the N-terminus of PAN might play an important role in regulating nucleotide hydrolysis, substrate protein recognition, and/or other activities. The MjPAN(Δ 1-73) complex still possesses both the highly conserved Walker A binding motif for the β and γ phosphates of nucleotides (208, 271) and the Walker B motif involved in Mg^{2+} -binding and ATP hydrolysis (27). Therefore, the MjPAN(Δ 1-73) protein provided the opportunity to examine the role of the highly charged N-terminus on the biochemical properties of the MjPAN protein. The ATPase activity of the full-length MjPAN complex is highest at 80°C which is close to the growth temperature optimum of 83° C for this organism (Fig. 7A). However, the truncated MjPAN(Δ 1-73) complex was relatively less heat stable, with an optimal ATPase activity at 65°C and a loss of activity at higher temperatures (Fig. 7A). While this study did not test the temperature optimum of the complex formed of both MjPAN and MjPAN(Δ 1-73), the ATPase activity of the complex of His₆-MjPAN and MjPAN(Δ 1-73) was recently shown to be optimal at 73°C (303). Secondary structural analysis suggests that an α -helix in the N-terminus of MjPAN, which contains a putative coiled-coil region (196, 275), is absent in the truncated protein. Therefore, the α helix may play an important role in stabilizing MjPAN at extreme temperatures. Based on analogy to other proteins that are stable at extreme temperature and/or salt conditions, it is possible that ion-pairs, hydrophobic clusters, or other structural elements are required

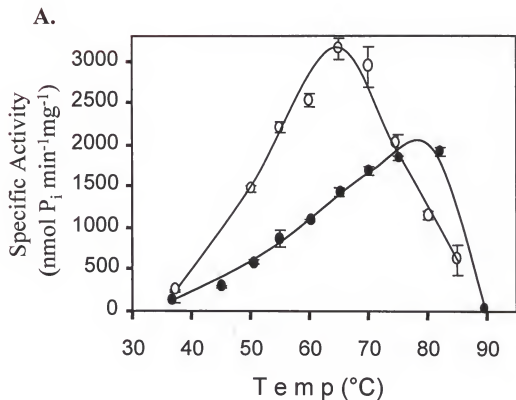
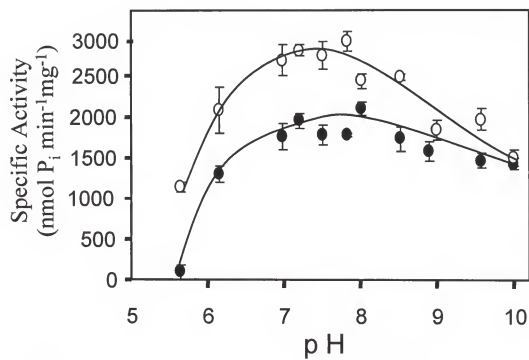
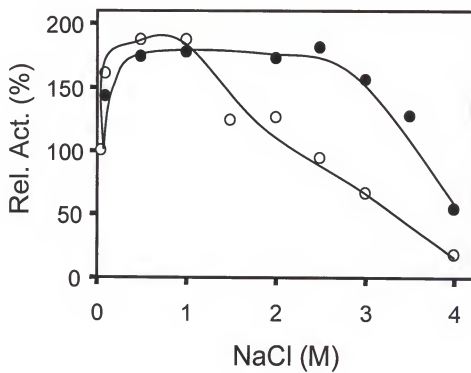


FIG. 7. Effect of pH, temperature, and [NaCl] on the ATPase activity of the MjPAN(●) and truncated MjPAN(Δ 1-73) (○) proteins. ATPase activity was measured as the release of inorganic phosphate at the temperature (A), pH (B), and salt concentrations (C) indicated.

B.**C.**

for the thermostability of MjPAN and that these are disrupted by removal of the N-terminal α -helix.

Both the full-length and truncated MjPAN proteins have optimal ATPase activity at pH 7-8. Both were also more active at higher pH (8-10) than at lower pH (below 6) (Fig. 7B). The ATPase activities of the full-length and truncated MjPAN proteins were both stimulated by the addition of either KCl or NaCl (Fig. 7C), and the proteins lost activity when preincubated in a salt-free buffer prior to the assay. The ATPase activity of full-length MjPAN was stimulated up to two-fold by NaCl or KCl concentrations to 2.5 M, while activity was reduced at higher salt concentrations. In contrast, the MjPAN(Δ 1-73) complex had lower activity in NaCl or KCl concentrations above 1M (Fig. 7C). In summary, the full-length MjPAN appeared to be more stable at both high salt and high temperature conditions. These results indicate that the full-length MjPAN protein, which was also detected in *M. jannaschii* using Western blot (Fig. 2), is likely the native form synthesized by this organism.

The ATPase activities of both PAN proteins were inhibited by the Mg^{2+} -chelator EDTA and also by the sulfhydryl-blocking agents NEM and *para*-chloromercuriphenylsulfonic acid (PCMS) (Table 5). The MjPAN activities were not significantly affected by NaN_3 (Table 5) which inhibits H^+ -transporting, membrane-bound ATPase proteins. These results are consistent with results obtained from studies of other proteins of the AAA family which catalyze Mg^{2+} -dependent ATP hydrolysis and are inhibited by NEM (11, 57, 98, 153, 241, 289). It has been proposed that a highly conserved cysteine residue in the C-terminus of these AAA-ATPases (*e.g.*, Cys384 of MjPAN) is most likely

TABLE 5. Effect of various compounds on the ATPase activity of the *Methanococcus jannaschii* PAN and truncated PAN ($\Delta 1-73$) proteins.

Effector	Conc. (mM)	Relative Activity ^a	
		MjPAN	MjPAN ($\Delta 1-73$)
None	-	100 $\pm$ 6	100 $\pm$ 5
EDTA	10	8 $\pm$ 3	UD ^b
SDS	0.1	63 $\pm$ 5	nd ^c
	0.2	3 $\pm$ 4	nd
NEM	10	UD	15 $\pm$ 4
	5	23 $\pm$ 1	nd
	1	81 $\pm$ 8	nd
PCMS	0.05	UD	UD
NaN ₃	5	90 $\pm$ 2	74 $\pm$ 5

^aATPase activity is expressed as a percentage of the control reaction with no added effector. The enzyme (2 μ g per ml) was preincubated with or without the effector in 100mM NaCl, 10mM MgCl₂, 25mM TES, pH8.0 at 21°C for 15 min before addition of 1mM ATP and incubation at 80°C (MjPAN) or 65°C (MjPAN $\Delta 1-73$).

^bUD, undetectable.

^cnd, not determined.

modified by PCMS and NEM (294). However, unlike other related AAA proteins such as Cdc48 of *S. cerevisiae* (57), preincubation of MjPAN with ATP did not protect the protein from NEM inhibition (Fig. 5, Lanes 7 and 8).

The purified MjPAN protein catalyzed a very high rate of ATP hydrolysis when compared to other AAA family members (Table 6) with a $V_{\max}$ of 3,500 nmol P_i released $\cdot \text{min}^{-1} \cdot \text{mg protein}^{-1}$ (Table 7). The only comparable ATPases are *E. coli* ClpA (104) and the yeast Cdc48 complex (57) with $V_{\max}$ values of ~5,000 and 1,560, respectively (Table 6). Several complexes which contain AAA-ATPases, such as the 19S complex of the 26S proteasome (98), have over 100-fold lower $V_{\max}$ values for ATP hydrolysis than MjPAN (Table 6). The affinity of MjPAN for ATP ($K_m = 497 \mu\text{M}$) is comparable to some other purified AAA proteins including yeast Cdc48 (57), mammalian NSF (165), and *E. coli* ClpX (278) (Table 6). However, MjPAN had a much lower affinity for ATP than some other AAA proteins including *E. coli* proteins ClpB (290), Lon (280), and FtsH (254). Also, the Rpt2p subunit as well as the 19S cap of the 26S proteasome have a much higher affinity for ATP than does MjPAN (98, 153). In general, there appears to be a wide range of both ATPase activities as well as affinities for ATP (at least *in vitro*), even among very closely-related AAA homologs.

Surprisingly, the $V_{\max}$ values for CTP hydrolysis by MjPAN and MjPAN(Δ 1-73) were 1.7- and 2-fold higher, respectively, than the $V_{\max}$ values for ATP hydrolysis (Table 7). This is consistent with the partially purified MjPAN complex from *M. jannaschii* cell lysate which had a 2:1 ratio of CTP- to ATP-hydrolyzing activity and was also inhibited by NEM. Full-length MjPAN had a 60% higher affinity for CTP than for ATP, while the truncated MjPAN(Δ 1-73) complex showed little change in affinity between the two

TABLE 6. Comparison of $V_{\max}$ and K_m values for AAA-ATPases.

Protein Name	Organism	$V_{\max}$ (nmol min ⁻¹ mg ⁻¹)	K_m (μ M ATP)	Reference
MjPAN	<i>M. jannaschii</i>	3500	497	this study
MjPAN (Δ 1-73)	<i>M. jannaschii</i>	4800	307	this study
Cdc48	<i>S. cerevisiae</i>	1560	550	(57)
p97	<i>X. laevis</i>	720	nd	(189)
VAT	<i>T. acidophilum</i>	90	nd	(181)
ARC	<i>R. erythropolis</i>	268	200	(289)
NSF	Hamster ovary	67	650	(241)
	(AAA-D2)	6	110	(165)
	(AAA-D1)	257	43300	(165)
Vps4p	<i>S. cerevisiae</i>	448	nd	(11)
Vhs4p (Rpt2p)	<i>S. cerevisiae</i>	7	5	(153)
Vhs4p Δ (Rpt2p Δ)	<i>S. cerevisiae</i>	20	5	(153)
SUG1 (Rpt6p)	Rat liver	7	35	(155)
26S proteasome	Rabbit reticulocytes	19	15	(98)
19S (RP)	Rabbit reticulocytes	34	30	(98)
ClpA	<i>E. coli</i>	5000	210	(104)
ClpB	<i>E. coli</i>	50-100	1100	(290)
ClpX	<i>E. coli</i>	500	500	(278)
FtsH	<i>E. coli</i>	230-459	80-83	(254, 111)
HslU (ClpY)	<i>E. coli</i>	57 (-HslV); 213 (+HslV)	280	(295)

nd, not determined

TABLE 7. Kinetic parameters of the *Methanococcus jannaschii* PAN and truncated PAN ($\Delta 1-73$) proteins.

Enzyme	Substrate	$K_m^{a,c}$	$V_{max}^{b,c}$
MjPAN	ATP	497	3500
	CTP	307	5800
MjPAN ($\Delta 1-73$)	ATP	477	4800
	CTP	493	9500

^a μ M of ATP or CTP.

^bnmol P_i released per min per mg protein.

^cFor determination of the kinetic parameters of MjPAN and MjPAN($\Delta 1-73$), enzyme (4 μ g per ml) was incubated with 1mM ATP or CTP in 25mM TES buffer at pH8.0 with 100mM NaCl and 10mM $MgCl_2$ at 80°C for PAN or 65°C for MjPAN($\Delta 1-73$). Initial velocity was determined from 0 to 6 min by measuring the release of P_i in triplicate. Initial velocity was plotted as a function of substrate concentration using Lineweaver-Burk, Michaelis-Menton, and Eadie-Hofstee plots (39).

nucleotides (Table 7). Of the nucleotides tested, the full-length MjPAN protein hydrolyzed only ATP and CTP and did not hydrolyze UTP, TTP, GTP, or ITP (Table 8). In contrast, MjPAN(Δ 1-73) hydrolyzed ITP, GTP, TTP, and UTP at 92, 73, 37 and 19% the rate of ATP hydrolysis, respectively (Table 8). Neither MjPAN or MjPAN(Δ 1-73) appeared to hydrolyze nucleotide diphosphates ADP or CDP (Table 8). The ATPase activity of full-length MjPAN was strongly inhibited by equimolar concentrations of ADP to 5% of the original activity but was only slightly inhibited in the presence of an equimolar concentration of CDP (Table 9). In contrast, the CTPase activity of full-length MjPAN was almost completely inhibited by ADP and also strongly inhibited by CDP to 16% of the original activity. For the truncated MjPAN(Δ 1-73) complex, the ATPase activity was also strongly inhibited by equimolar ratios of ADP to 10% of the original activity, while CDP had no detectable effect on ATPase activity. The CTPase activity of MjPAN(Δ 1-73) was not as strongly inhibited by either ADP or CDP as was full-length MjPAN (Table 9). Inhibition by ADP has also been observed for other AAA-ATPases including ClpB (290) and HslU (201). Although the high rate of CTP hydrolysis was somewhat surprising, the results are consistent with previous studies which have shown that CTP can substitute for ATP hydrolysis in assembly of the eukaryotic 20S proteasome with the RP complex (60) containing the six AAA Rpt proteins of the Base domain (72). In addition, another recent study of the MjPAN complex reported the hydrolysis of GTP and UTP at 58% and CTP at 205% the rate of ATP hydrolysis (303). Other AAA-ATPases have also been observed to hydrolyze several types of nucleotides, including the 19S cap of the mammalian 26S proteasome and *E. coli* Lon which can hydrolyze ATP, CTP, UTP, and GTP (98, 279). The ARC complex from *R. erythropolis* (289) can

TABLE 8. Nucleotide-hydrolyzing activities of the *Methanococcus jannaschii* PAN and truncated PAN(Δ 1-73) proteins.

Substrate	Relative Activity ^a	
	MjPAN	MjPAN (Δ 1-73)
ATP	100 $\pm$ 7	100 $\pm$ 4
CTP	189 $\pm$ 5	209 $\pm$ 8
UTP	UD	19 $\pm$ 6
GTP	UD	73 $\pm$ 3
TTP	UD	37 $\pm$ 0.3
ITP	UD	92 $\pm$ 1
ADP	UD	UD
CDP	UD	UD
AMP-PNP	UD	UD

^aActivity is expressed as a percentage of inorganic phosphate released compared to the control reaction using the indicated substrate. Enzyme (2 μ g per ml) was incubated with 1mM nucleotide in 25mM TES buffer at pH 8.0 with 100mM NaCl and 10mM MgCl₂ at 80°C for MjPAN or 65°C for MjPAN(Δ 1-73).

UD, undetectable

TABLE 9. The effect of dinucleotides on nucleotide triphosphate hydrolyzing activities of recombinant *M. jannaschii* PAN and PAN ($\Delta 1-73$) proteins

	Relative Activity ^a	
	MjPAN	MjPAN ($\Delta 1-73$)
<u>CTP Hydrolysis</u>		
Control	100 $\pm$ 0.4	100 $\pm$ 2
ADP	UD ^b	30 $\pm$ 1
CDP	16 $\pm$ 0.6	23 $\pm$ 2
<u>ATP Hydrolysis</u>		
Control	100 $\pm$ 7	100 $\pm$ 4
ADP	5 $\pm$ 5	11 $\pm$ 3
CDP	87 $\pm$ 2	100 $\pm$ 0.8

^aRelative activity is expressed as the percentage of inorganic phosphate released compared to control reaction without added nucleotide diphosphate. The enzyme (2 μ g per ml) was preincubated with or without 1mM CDP or ADP for 15 min in 25 mM TES at pH 8.0 with 100 mM NaCl and 10 mM MgCl₂ at 21°C before addition of 1mM ATP or CTP and incubation at 80°C for MjPAN or 65°C for MjPAN ($\Delta 1-73$).

^bUD, undetectable.

hydrolyze both CTP and ATP at fairly high rates. The FtsH protease from *E. coli* hydrolyzes CTP, GTP, and UTP at 107, 46, and 27% the rate of ATP hydrolysis, respectively; however, only ATP or CTP supported degradation of the substrate σ^{32} protein (111).

PAN-Stimulated ATP-dependent Proteolysis of Protein Substrates
by the 20S Proteasome

Both full-length MjPAN and MjPAN(Δ 1–73) complexes stimulated the degradation of β -casein by the *M. jannaschii* 20S proteasome by 2.2 ± 0.19 -fold and 1.4 ± 0.04 -fold, respectively. This stimulation was completely dependent on the presence of either ATP or CTP (data not shown). No significant energy-dependent hydrolysis of β -casein was observed in the presence of ADP, and stimulation was similar at molar ratios ranging from 0.5:1 to 4:1 of PAN:20S proteasome complex. Furthermore, the MjPAN and MjPAN(Δ 1–73) complexes stimulated by 2- to 4-fold the degradation of β -casein by the *M. thermophila* 20S proteasome in the presence of ATP but not ADP (Fig. 8). Under these assay conditions, the rate of ATP hydrolysis by MjPAN was not influenced by the addition of the synthetic peptide Suc-LLVY-AMC or the 20S proteasome alone but was enhanced by the addition of β -casein by 1.22 ± 0.07 -fold or β -casein plus the 20S proteasome by 1.38 ± 0.08 -fold (data not shown). In addition, CTP could substitute for ATP (at ~80% the rate) in stimulation of protein hydrolysis (Fig. 8). This is consistent with what has been observed for other energy-dependent proteases including the eukaryotic 26S proteasome (60), as well as *E. coli* proteins Lon (279) and FtsH (254). Somewhat surprisingly, 5'-adenylyl β , γ -imidophosphate (AMP-PNP), which is not hydrolyzed by the MjPAN protein, also supported MjPAN activation of β -casein

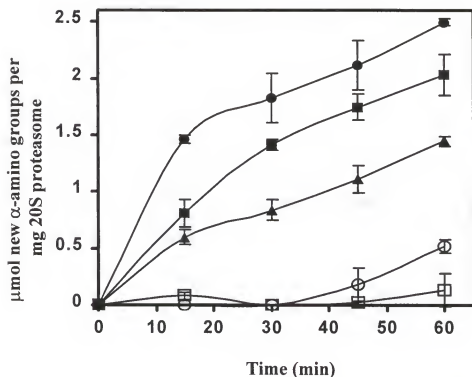


FIG. 8. Nucleotide-dependent MjPAN and MjPAN ($\Delta 1-73$) stimulation of β -casein degradation by the *Methanosarcina thermophila* 20S proteasome. The 20S proteasome was incubated at 55°C with or without addition of MjPAN proteins and/or nucleotides, as indicated. The buffer used was 25 mM TES (pH8) with 10mM $MgCl_2$, 1mM nucleotide, and 2 mg of β -casein per ml. Newly released α -amino groups generated by protein hydrolysis were measured using fluorescamine. Symbols: ●, MjPAN and 20S proteasome with 1mM ATP; ■, MjPAN and 20S proteasome with 1mM CTP; ▲, MjPAN ($\Delta 1-73$) and 20S proteasome with ATP; □, 20S proteasome alone with 1mM ATP; ○, MjPAN and 20S proteasome with ADP.

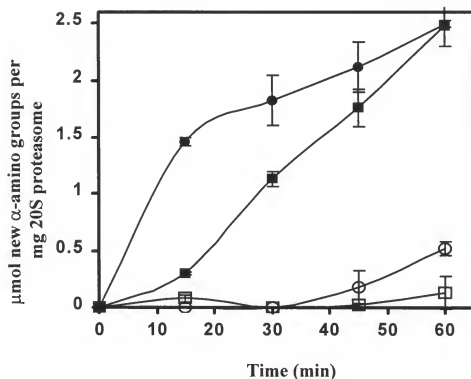


FIG. 9. ATP- and AMP-PNP-mediated MjPAN stimulation of β -casein degradation by the *Methanosarcina thermophila* 20S proteasome. The 20S proteasome was incubated at 55° C with or without addition of MjPAN protein and/or nucleotides, as indicated. The buffer used was 25mM TES (pH8) with 10mM $MgCl_2$, 1mM nucleotide, and 2 mg of β -casein per ml. Newly released α -amino groups generated by protein hydrolysis were measured using fluorescamine. Symbols: ●, MjPAN and 20S proteasome with 1mM ATP; ■, MjPAN and 20S proteasome with AMP-PNP; ○, 20S proteasome alone with ATP; □, MjPAN and 20S proteasome with ADP.

degradation (Fig. 9). However, the mechanism of degradation appeared to be qualitatively different. When ATP or CTP was hydrolyzed there was a rapid increase in the generation of free α -amino groups which then leveled off. This plateau in activity could be due to an increase in the ADP or CDP concentrations which then inhibit the nucleotidase activity of the MjPAN complex. In contrast, there was a significant lag phase observed when AMP-PNP was present. In addition, there was no indication that the generation of new α -amino groups was reaching a plateau (Fig. 9). Although these results contrast with those of another recent study which reported that the complex of His₆-PAN/PAN(Δ 1-73) in the presence of AMP-PNP does not activate the breakdown of proteins by the *T. acidophilum* 20S proteasome (303), the methods used to monitor protein hydrolysis were different. In the present study, hydrolysis was measured by the production of new α -amino groups using fluorescamine, whereas the previous study measured conversion of β -[¹⁴C]casein into acid-soluble products (303). Together, these results suggest that the presence of a nonhydrolyzable ATP analogue, such as AMP-PNP, may support only a single cleavage or a small number of cleavages of β -casein by the 20S proteasome. This would result in the generation of large, acid-precipitable fragments with new amino groups. Thus, it is possible that ATP hydrolysis is required for the MjPAN-stimulated processive degradation of proteins to acid-soluble products by the 20S proteasome. However, ATP (or a non-hydrolyzable analog such as AMP-PNP) binding to MjPAN supports limited and non-processive hydrolysis of proteins by the 20S proteasome. Similar results were seen with the Lon protease of *E. coli* (52). Edmunds and Goldberg demonstrated that non-hydrolyzable nucleotide analogs such as AMP-PNP could support cleavage of casein by Lon into several large fragments, but processive

degradation into acid-soluble small fragments was only observed in the presence of ATP (52). The HslUV protease can also support proteolysis of protein and peptidase substrates in the presence of AMP-PNP, but there is an initial lag time in proteolysis that is not seen with ATP (103). Interestingly, HslU has two C-terminal cysteine residues, and mutation of one (C261V) blocks the NEM-mediated monomerization of the complex, but has no effect on ATPase activity (294). Mutation of the other C-terminal cysteine residue (C287V) abolishes the ATPase activity of the complex; however, the mutant HslU can still form a complex with HslV to support protein hydrolysis (294). These results indicate that ATP binding, but not necessarily hydrolysis, is required for oligomerization of HslU and for interaction of HslU with HslV. Finally, the nonhydrolyzable analog ATP- γ S can support the chaperone activity of ClpX in the protection of λ O from aggregation, but ATP hydrolysis is required for the disassembly of λ O aggregates (278).

The N-terminal coiled-coil region of the MjPAN protein. The effect of an N-terminal truncation has been studied in other AAA proteins, including ClpA (217) and ClpB (290) from *E. coli* and Vhs4p (Rpt2p) from *S. cerevisiae* (153). When the ClpB protein from *E. coli* was expressed from a plasmid, it purified as a complex composed of both 93- and 79-kDa proteins (236, 290). When the open reading frame is modified to produce only the 79-kDa protein from the internal-translation start site, a tetrameric complex still formed (183). The truncated ClpB complex had higher ATPase activity than the full-length complex, but the ATPase activity was not stimulated by protein substrates indicating that the N-terminal domain may contain a region essential for substrate protein binding (183). Furthermore, when increasing concentrations of the truncated ClpB were mixed with the full-length ClpB, the truncated ClpB was found to inhibit the protein-

activated (but not the basal) ATPase activity of full-length ClpB (183). Interestingly, an N-terminal deletion ($\Delta 1-50$) of the yeast Yhs4p also results in a 2.8-fold increase in ATPase activity compared to the full-length protein (Table 6); however, the ability of this modified protein to support ATP-dependent protein degradation has not been determined (153). Finally, based on a study where the human Rpt2p/S4 of the RP was synthesized without the N-terminal 85-amino-acid coiled-coil region, it was determined that this region of the protein is necessary for interaction with the Rpt1p/S7 subunit (196). These studies demonstrate that the shorter proteins derived from the internal translation sites of these ORFs act to inhibit the activity of the full-length protein. Similar results were obtained in this study. Even though the MjPAN($\Delta 1-73$) complex had higher ATPase activity than the full-length MjPAN complex, it supported the ATP-dependent hydrolysis of casein by the 20S proteasome to only ~60% the rate of MjPAN (Fig. 8).

Association of MjPAN with Archaeal 20S Proteasomes

The purified *M. jannaschii* 20S proteasome and MjPAN proteins were mixed at equimolar ratios and then incubated in the presence of ATP and Mg^{2+} at 50° C. Electron micrographs of these mixtures reveal particles which resemble the 20S core of 12 nm width by 17 nm length capped either singly or doubly on each end by MjPAN protein complexes (Fig. 10). The singly capped proteasome-associated MjPAN complexes were ~24 nm long and 12 to 14 nm wide. These micrographs are consistent with the ability of the MjPAN proteins to stimulate proteolysis by 20S proteasomes and suggest that archaeal 20S proteasomes may interact with MjPAN to form a complex which resembles the 26S proteasome subcomplex recently described in *S. cerevisiae* (72). The efficiency

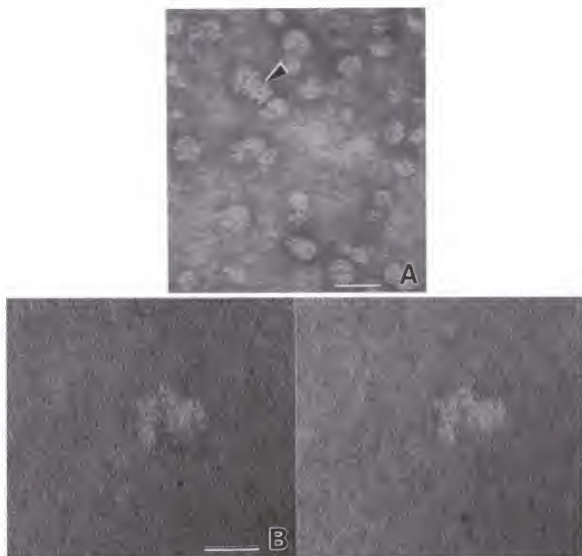


FIG. 10. Transmission electron micrographs of reconstituted 20S proteasome and MjPAN proteins. A. Untilted view of particle generated during reconstitution of the MjPAN($\Delta 1-73$) and 20S proteasome. Arrowhead indicates cylindrical assembly with apparent comma-shaped assemblies of putative MjPAN($\Delta 1-73$) dodecamers at both ends. B. Stereo view of single 20S cylinder with putative comma-shaped MjPAN assembly at left end. Bars, 20 nm.

of reconstitution of the MjPAN:20S proteasome, however, was very low (less than 0.5%), indicating that either 1) the interaction between the proteins may be transitory, 2) the complexes are not stable during preparation for electron microscopy, or 3) additional factors may be necessary for stable complex formation.

Proteasomal Components in *M. maripaludis*

Even though the entire genome of *M. jannaschii* has been sequenced, and several recombinant proteins have been characterized, little is known of the physiology or genetics of this hyperthermophilic archaeon. While *in vitro* studies using recombinant proteins have been useful, a better understanding of the physiological functions and substrates of the energy-dependent proteolytic system in methanogens and other archaea will require the ability to perform *in vivo* studies. Both genetic tools and high-yield culture techniques have been developed for use in the mesophilic methanoarchaeon *M. maripaludis*. Over-expression of recombinant proteins which require the reducing environment provided by the *M. maripaludis* cytosol, and thus are not expressed in an active form in *E. coli*, has also been demonstrated (62). These techniques should aid in purification and characterization of recombinant proteins from methanogens.

Isolation of the Mmp_{pan} Gene from a *M. maripaludis* Genomic Library Southern analysis was performed using *M. maripaludis* genomic DNA which had been digested with several restriction enzymes, including *EcoRV*, *XbaI*, and *SmaI*. When the Southern blot was probed using the *M. jannaschii pan* ($\Delta 1-73$) gene fragment (Mj_{pan} probe), a 9-kbp *XbaI* and a 2.5-kbp *EcoRV* positive fragment were identified. Digestion of genomic DNA with *SspI* (with recognition sequence AAT⁺ATT) yielded numerous small positive fragments, due to the high A-T content (67%) of the *M. maripaludis* genome. A random

M. maripaludis genomic λ library was then screened using the Mjpan probe to isolate the related gene from *M. maripaludis*. Approximately 1600 total plaques were screened, and 14 positive clones were identified. Plasmid subclones from all 14 positive clones were then obtained by automatic Cre-lox-mediated subcloning using the host strain BM25.8 (Table 3). This procedure yielded >500 colonies for all 14 clones after incubation at 37°C overnight, and three clones (#3, #8, and #11) were chosen for further study. Clone #11 contained a plasmid (pJAM 721, Table 4) with a ~15-kbp insert, which yielded a 9-kbp *Xba*I and a 2.5-kbp *Eco*RV positive fragment during Southern analysis. The 407 amino-acid long deduced sequence of the isolated 1224-bp ORF (Fig. 11) was 70% identical to MjPAN. Interestingly, when the ORF was examined, two possible translation start sites were detected, both with putative Shine-Dalgarno consensus sequences (Fig. 11). However, there was an in-frame stop codon 39-bp downstream from the first start site, suggesting that translation from this site would produce a 16 amino-acid long peptide. To confirm the accuracy of the DNA sequence and to ensure that the correct start site was chosen for cloning of Mmp_{pan}, the upstream region 5' to the start site was amplified directly from *M. maripaludis* genomic DNA using the high-fidelity Vent DNA polymerase. The PCR fragment was cloned into a pCR-TOPO vector to create plasmid pJAM745 (Table 4). Sequencing of this plasmid in both directions verified that the upstream region of the gene isolated directly from genomic DNA was identical to that obtained from the λ library clone. In addition, when the Mmp_{pan} sequence from this study was compared to Mmp_{pan} (ORF558) from the *M. maripaludis* LL genomic sequence (J. Leigh, personal communication), the DNA sequence was 96% identical. Furthermore, the nucleotide substitutions were all

5'-cttggttttgttttgtttcatttttgaatactgccttttggttttctagaatatcctt
taggcgacattccaaattttgcgaacgctcgcaaacattcatttctacgccttctactc

Oligo #1: 5'-**cctctttttccgcagagttcgcat**tg-3'
ttgttttaatgatatttcgtaacctcttttccgcagagttcgcatcgtacacacacctt
aatatactaaatttaagatatttttgcgaagtatcctgattacatagcattcaaatg
gtaaatgcttcattagagaatatatcttcaattatttaatatattctatttttgcatttt

Box A Box B
atatattatatatcc**ataaaatttatata**gtccaaataaaattataaaacat**atac**ccttattgga
ataatagttgttaaaaatgttttcgataggtgtaaaaactgttttgcgagaaagaaaatca
atgagaaagcgctatatagaactatttttatcatatttattacggagg**taa**at**cat**gagt
M R K R Y I E L F L S Y L L R R * I M S
taccagatgattattccacagatatagaaaaataaaatggatttaaagaatttaaa
Y P D D Y S T D I E K N K M D L K E F K 22
gaaaaaactcagatcgacagtttagaaagcaaagtattgagattggaattaaaaataaa
E K T Q I A E L E S K V L R L E A L K N K 42
gatattaatagagagaatgtacaaattaaaaaagaaaatgaaatattaaagcgagaactt
D I N R E N V Q I K K E N E I L K R E L 62
gataaattaagaattccaccactcatactcgggtacaatacttgataaagtaaataagaa
D K L R I P P L I L G T I L D K V N E R 82
aaggctgtgttaaaagtcttacagggccaaaacttcttgtaaaccctttcacattcgta
K A V V K S S T G P N F L V N L S Q F V 102
gatcctgaggatattgttcacagagcaagagtttgtttaaacacagcagacttttagcgatt
D P E D I V P G A R V C L N Q Q T L A I 122

Oligo #2: 3'-**gtggctcgataccggttac**ctttag-5'
gttgaagtattgcctaaagaaaaagactaccgagctatggcaatggaaatcgaggaaaaa
V E V L P K E K D Y R A M A M E I E E K 142
ccagacatttcccttgaagatatcggtggtttgaacaaccagattagagatatataagaa
P D I S F E D I G G L N N Q I R D I K E 162
gtagtagaactcccacttaaaaaccagaattatttgaaaaagtgggtatcggtccacca
V V E L P L K N P E L F E K V G I V P P 182
aaaggagtattactctatggtccgcaggaaactggaaaaacgctccttgcaaaagctggt
K G V L L Y G P P G T G K T L L A K A V 202
gcatacgaaaccaatgcatcatttgttagagttgtcgctcagaactcgtaaaaaatttc
A Y E T N A S F V R V V G S E L V K K F 222

FIG. 11. Deduced amino acid sequence of the *Mmp*an gene. Putative ribosomal binding sites are double underlined, putative start codons are in bold, and putative stop codons are indicated by an asterisk. Amino acid positions of the MmPAN protein deduced from the DNA sequence are indicated on the right. Oligonucleotides used to generate plasmid pJAM 745 are indicated above the nucleotide sequence. Putative promoter sequence (Box A) and transcription start site (Box B) are indicated.

```

atcggtggaaggtgcaaaacttgtaagagatgttttcaaacttgcaaaagaaaaatcccc
I G E G A K L V R D V F K L A K E K S P 242
tgcatacatatttattgatgaaatcgatgcagttgcaagcaaaagaactgaatcattaacc
C I I F I D E I D A V A S K R T E S L T 262
ggtggagatagagaagttcaaagaactttaatgcagcttcttgcagaaatggatggtttt
G G D R E V Q R T L M Q L L A E M D G F 282
gattcaagaggcgatgtaaaaaatcattgctgcgacaaacagaccagacattttgatcct
D S R G D V K I I A A T N R P D I L D P 302
gcaatactcagacctggaagatttgatagaatcattgaaatttcaatgcctgatgaagat
A I L R P G R F D R I I E I S M P D E D 322
ggaagactcgaaatcttaaaaaatccacactgaaaaaatgaaccttaaagggtgtggactta
G R L E I L K I H T E K M N L K G V D L 342
agagaagttgcaaaaattgcagaaaacatggctcgggtgctgacttaaagcagtatgtact
R E V A K I A E N M V G A D L K A V C T 362
gaagctgggtatgtttgcaataagagaagaaagagaattcatcaaaatggatgacttcaga
E A G M F A I R E E R E F I K M D D F R 382
gaagctatctcaaaaatcacaggtaaagaaagaaaaatgcagttacgacatgccacaatta
E A I S K I T G K K E K C S Y D M P Q L 402
accgtaatgtacgggataaattaactactattttttcaacttatttttatttttaaaattttt
T V M Y G * 407
taagtgatgtttatgaagattggaattattggagcaggccttgggggcctattaagcggt
gcaattttatcaaaagaacactcagttacaatttacgaaaaacttcatttgggtggaggt
agattacaaatattcctacaaaggattcaactgacaacaggagccttcacatgatcctca
tggaactacgggtatttagcaca -3'

```

FIG. 11. Cont.

conservative, and the deduced amino acid sequences of the two ORFs were identical. The ORF558 also had an identical 5' region, with conserved in-frame start and stop codons as well as two putative ribosomal binding sites. Thus, both *M. maripaludis* strains may produce an identical small peptide of 16 amino acids with the sequence ***MRKRYIELFLSYLLRR*** (basic and hydrophobic residues are indicated by italics and bold, respectively). This putative peptide is highly basic at the N- and C-termini, with patches of hydrophobic and aromatic residues in the middle. The N-terminus of MjPAN is also highly charged with alternating basic and acidic residues but very few hydrophobic or aromatic residues. Whether or not this small peptide is produced in the cell and what purpose it might serve has not been determined at this time. As can be seen from a multiple sequence alignment of archaeal PAN proteins (Fig. 12A), MjPAN has an N-terminal extension that is not seen in any other archaeal PAN homolog. This included PAN proteins from other hyperthermophilic organisms. A sequence alignment of the N-terminal region of MjPAN to the 16 amino-acid long peptide upstream from MmPAN reveals that the N-terminal region of MjPAN has some identity to this peptide (Fig. 12B). Biochemical results from this study (discussed above) indicate that the N-terminus of MjPAN may help stabilize the protein in high temperature. Thus, it is possible that the N-terminus of MjPAN, which is required for the extreme thermostability of this protein in *M. jannaschii* is dispensable in the mesophilic *M. maripaludis*. A future experiment to test this hypothesis would involve mutating the stop codon of this peptide so that MmPAN is produced with an 18 amino-acid long N-terminal extension. The thermostability and other biochemical properties of this protein could then be compared to the wild-type MmPAN.

A.

```

MmPAN      1 -----MSYPDDYSTDIEKNKMDLKEFKETQIAEIESKV
MjPAN      1 ---MVFEFISTELKKEKAFTEEFKEEKEINDNSNNDLLKEELQEKARIAEIESH
AF1976     1 -----MGDSEIQYLLEKLKLEEDYIKRRLY
MTH0728    1 -----MENNSQNVLKIEDIKKEVRLMKENSKTGRNMNRI
PH0201     1 -----MSVSGDEVQFGDYYDDIITYKKRI
PAB2233    1 -----MNIISGDEVQFHENYDDIITYKKRI
APE2012    1 -----MTISAGGGRSHRHNGHGHSSERDVEI
AAG20171   1 -----MSRSPSLPDRPTLDVDPSTPAERLNALQDHVYDIAVAN
AAG19042   1 -----

```

```

MmPAN      35 LRLELNKQDINRENVQIKKENEILKRELDKLRIPPLILCTILDVNERKAVVKSSTGPNF
MjPAN      57 LKLELEKKELEBRENLOIMKENEILRRELDRMVPLPIVGTVDVKVGRKRVVKSSTGPNF
AF1976     28 RRLDEKKEFIESERIPYEREVRLRSEVERILRSPPLIVGVMSDILEDGRNVVKSSTGPNF
MTH0728    38 RRLKDKLLIENETRLDREVKSLRGEIERERAPPLVIATVEVLDDHRAVAVESTGPNF
PH0201     27 ROLELOVRMLPADKRLERELSLRSENRLRPPFAFAGTVIEVLDDHRAIVQNYNGRPF
PAB2233    27 ROLELOVRMLPADKRLERELSLRSENRLRPPFAFAGTVIEVLDDHRAIVQNYNGRPF
APE2012    26 RRLKDKVRSLSLTKKISLQKELEYEYNEITKLLSPPIHAWVLEIDDNRVVKSSTGPNL
AAG20171   40 GELCAQLDDVBARREEDREYVRLQRENEILKTASLYLATVGLDPEIGSAVIKCHGNQGE
AAG19042   1 --YDRRLDANAENNYQOKLRLSHENKRLQSPSTIATVGLDPEIG--AVIKCHGNQGE

```

```

MmPAN      95 LVNLSQFVDPDITVPGARVCLNQOTLAIVEVLPKEKDYRAMAMEIEEKPDISEDIGGN
MjPAN      117 LVNVSHENPDILAPGRVCLNQOTLTVDVLPENKIDYRAKAMEVDERPNVVRKEDIGGLE
AF1976     88 VVNTSQYNBEBELPPGARVALNQOTLAIVNVLPSTKDPVYGEFVEEKPEVVEDIGGLD
MTH0728    98 VVNTSREIDRKLPPGARVALNQOTIESIVDVLPSKDDPVVTEMEVEEKPDVSYEDIGGLE
PH0201     87 VVRILAPWIDRKLPPGARVALDQRTMAVEILPSTKDPVILCFEVIERNVYNDIGGLK
PAB2233    87 VVRILAPWIDRKLPPGARVALDQRTMAVEILPSTKDPVILCFEVIERNVYNDIGGLK
APE2012    86 IVLVIAAGVILARSILPPGARVALNNGSTIVDVLPGRYDIPVYKAMEVEERPKVFEKDWGGL
AAG20171   100 VLTSLPRLADILEVGRVALINDPSVQVRLDDETARAQAMEVDSPSVTHADIGGLD
AAG19042   58 ALTEVTDELRLATLEPGSRVAVN--NSLSVVRQLDDEAVRVRVMEVDSPPVDIGEDIGGLD

```

```

MmPAN      155 NOIRDIKEVVELPLKNPELFEVGIVPPKGVLLYGPPGTGKTLAKAVAEATNASFVRVV
MjPAN      177 KQCEIREVVELPLKHPLEFVKVGIPEPKGILLYGPPGTGKTLAKAVAEATNATFIRVV
AF1976     148 VQREIREAVELPLKPELFAEVGIEPPKGVLLYGPPGTGKTLAKAVAEATNATFIRVV
MTH0728    158 EQVREVKTEVVELPLKHPLEFVKVGIPEPKGVLLYGPPGTGKTLAKAVAEATNATFIRVV
PH0201     147 KQCEIREATELPLKHPLEFVGIIDPPKGVLLYGPPGCGKTLAKALAHEVNTATFIRVV
PAB2233    147 KQCEIREATELPLKHPLEFVGIIDPPKGVLLYGPPGCGKTLAKALAHEVNTATFIRVV
APE2012    146 EQREIRYEAVVLPKNEHLEREGIDPPKGVLLYGPPGTGKTLAKAVAEATNATFIRVV
AAG20171   159 DQREVRVAVEDPIVNPKEIDVGVGPEPPSGVLLYGPPGTGKTLAKAVAEATNATFIRVV
AAG19042   117 DQREVRVAVELPMKPELFEVGIIDPPKGVLLYGPPGTGKTLAKAVAEATNATFIRVV

```

FIG. 12. A. Multiple sequence alignment of Archaeal PAN homologs. Abbreviations: AF1976, *A. fulgidus* PAN; MTH0728, *M. thermoautotrophicum* PAN; PH0201, *P. horikoshii* PAN; PAB2233, *P. abyssi* PAN; APE2012, *A. pernix* PAN; AAG20171 and AAG19042, *H. volcanii* PAN paralogs. Walker A and B consensus sequences are boxed. Conserved cysteine residue in C-terminus is indicated by arrows above and below the sequence. B. Pairwise sequence alignment of the N-terminus of MjPAN and the 16-residue ORF in the upstream region of MmPAN.

Walker B

```

MmPAN 215 GSELVKKF IGEGAKLVRDVFRLAKEKSPCLIFIDE IDAVASKRTSEH TGGDREVORTLMO
MjPAN 237 GSELVKKF IGEGASLVKDIFRLAKEKAPSIIFIDE IDATAAKRTDAL TGGDREVORTLMO
AF1976 208 GSEFVQKY IGEGARLVREVFOLAKEKAPSIIFIDE IDATAARRTNSD TSGDREVORTLMO
MTH0728 218 ASEFVVKY IGEGARLVRCVFRLAKEKSPSIIFIDE IDAVAAKRLKSTSGDREVORTLMO
PH0201 207 GSELVKKY IGEGARLVHELFE LAKEKAPTII FIDE IDAIGAKRDETTGGEREVNRTLMO
PAB2233 207 GSELVKKY IGEGARLVHELFE LAKEKAPTII FIDE IDAIGAKRDETTGGEREVNRTLMO
APE2012 206 GSELVKKY IGEGARLVREIFRLAKEKAPSIIFIDE IDATAKRVDTCTSGDREVORTLMO
AAG20171 219 GSELVKKF IGEGSRLVRDLFE LAEOKOPAIIFIDE IDAVAAKRTDSTSGDREVORTLMO
AAG19042 177 GSELVHKF IGEGAKLVRDLFQ VARDHEPAAVVFIDE IDATAKRTDSTSGDREVORTLMO

```

```

MmPAN 275 LLAEMDGFDSAGDVKI IAAATNRPDILDPAILRPGFRDRIETSPDDEGRLEILKIHTK
MjPAN 297 LLAEMDGFDAAGDVKI IGAATNRPDILDPAILRPGFRDRIEVPAPDEKGRLEILKIHTK
AF1976 268 LLAELDGFDPAGDVKVI GATNRDILDPAILRPGFRDRIEVPLETFEGRIQIEIKHTRK
MTH0728 278 LLAELDGFDSRCNVCI IAAATNRPDILDPAILRPGFRDRIEVPLENGGRLEILKIHTSG
PH0201 267 LLAEMDGFDPAGNVKVI IAAATNRPDILDPAILRPGFRDRIEVPLEDFEGRLEILKVHTRR
PAB2233 267 LLAEMDGFDPAGNVKVI IAAATNRPDILDPAILRPGFRDRIEVPLEDFEGRLEILKVHTRR
APE2012 266 LLAELDGFDPDNLNVKVI IAAATNRDILDPAVL RPGFRDRIEVPLESLRGRLEILGHTRK
AAG20171 279 LLAEMDGFDPGRGDIRI IAAATNRDMLDPAILRPGFRDRIEVPNPNPARERIIEITHAGE
AAG19042 237 LLAEMDGFDPGRGDIRI IAAATNRDMLDPAILRPGFRDRIEVPNPNVGGREKIFRIHTRA

```

▼

```

MmPAN 335 MNLKKG-VDLREVAKIENMV GADLKAVCTEAGMFAIREEREFLKDDFREATSKITTEKKE
MjPAN 357 MNLADVDNLEIAKMTGECVGAELKAICTEAGMNAIREDRDYVTMDDFRKAVEKIMEKKK
AF1976 328 MKLADVDVFELARITEGASGAD IKAICTEAGMFAIREERAKVTMLDTKAIKVKIKKTT
MTH0728 338 MNLAEVDVTEILARITD GASGADLKAICTEAGMFAIRDERDEVTMADEMDAVDKIMGVEK
PH0201 327 MKLKG-VDLRALAEMTEGASGADLKAICTEAGMFAIREERTYVTQEDFLKAVDKVLGNER
PAB2233 327 MKLKN-VDLRVLAETEGASGADLKAICTEAGMFAIREERTYVTQEDFLKAVDKVLGNER
APE2012 326 AKMAPDVLEIAKLTGEGSADLKAVVVEAGYNAIKRGSRVITMDMIKAVEKVKKALD
AAG20171 339 MNVADSVVFSLLAADTAERSGAQLASLATEAGMFAIRDDEVRHQEDDDEYKLVAEGD
AAG19042 297 MNVADSVVFGLLAADTSDL SGADVKAICTEAGMFAIRDDEVRHQEDDDEYKLVAEGD

```

▲

```

MmPAN 394 KCSYDMPQLTYMYG-----
MjPAN 417 VKVKEPAHLDTLYR-----
AF1976 388 PIPDLKGVMEVY-----
MTH0728 398 EEEYKQETGVYFG-----
PH0201 386 KLLQQITSHETIYG-----
PAB2233 386 KLLQQITSHETIYG-----
APE2012 386 KRGGGDPFIRAQKSGDDTIATVI
AAG20171 399 TESSGPRPSYTHQ-----
AAG19042 357 PAAATDVSRFTA-----

```

B.

```

MmPAN MRKRYIETELSYLRRIM
MjPAN --MVREEFISTELKKEK

```

FIG. 12. Cont.

Purification of MmPAN from Recombinant *E. coli*

The Mmpan ORF was cloned into pET24b and pET15b expression vectors to generate plasmids pJAM727, pJAM728, and pJAM730 (Table 4). These plasmids were used to synthesize MmPAN in recombinant *E. coli* strains supplemented with the rare tRNA genes. The MmPAN proteins that were produced in recombinant *E. coli* cross-reacted with the MjPAN antibody based on Western analysis (data not shown). The N-terminal His₆-tagged MmPAN protein (N-His₆-MmPAN) formed inclusion bodies and aggregates, with a small amount of protein in the soluble fraction. The soluble protein bound to a Ni²⁺-Sepharose column, but had no detectable ATPase activity and formed aggregates when stored at 4°C. The C-terminal His₆-tagged protein (C-His₆-MmPAN) formed fewer aggregates and inclusion bodies than N-His₆-MmPAN, and more protein was in the soluble fraction. The C-His₆-MmPAN fractions which eluted after purification using a Ni²⁺-Sepharose column were loaded onto a Superose 6 size exclusion column. The C-His₆-MmPAN eluted from the column in a broad peak from >1.5-MDa to 500-kDa, but very little ATPase activity was detected in these fractions. For purification of the unmodified MmPAN protein, *E. coli* (pJAM727) cell-free extract was applied to a Superose 6 column. While some MmPAN protein eluted in the void volume, (>1.5-MDa) there was also a peak of MmPAN at 600-kDa. These fractions were further purified using a MonoQ column, where the protein eluted at 500 mM NaCl. Based on non-denaturing native-gel electrophoresis (Fig. 13), this 600-kDa fraction formed a distinct complex which migrated similar to MjPAN. However, very little ATPase activity was detected for MmPAN, and the protein aggregated when incubated at 50° C (Fig. 13).

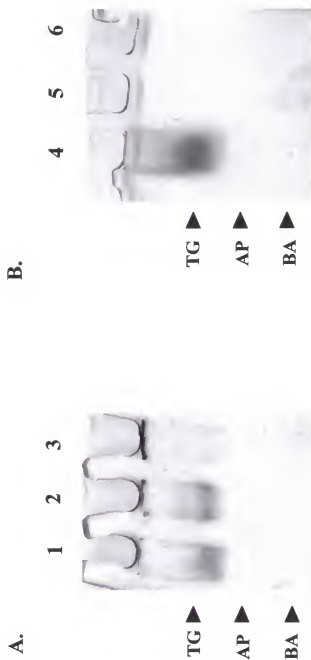


FIG. 13. Non-denaturing native gels electrophoresis of MmPAN. TG, thyroglobulin (669kDa). AP, apoferritin (443kDa). BA, β amylase (200kDa). A. Gel stained with Coomassie blue. B. In-gel ATPase activity was measured at 37°C. Lanes 1 and 4, Purified MjPAN (2 μ g); Lanes 2 and 5, Purified MmPAN (2 μ g); Lanes 3 and 6, Purified MmPAN (2 μ g) which was incubated at 50°C for 15 min before loading.

Partial Purification and Characterization of Two NTPases from *M. maripaludis*

During attempts to isolate MmPAN directly from *M. maripaludis*, two complexes of 200-kDa (NTPase #1) and 250-kDa (NTPase #2) with NEM-inhibitable nucleotidase activity were identified (Fig. 14) and partially purified (Fig.15). Neither of these NTPases cross-reacted with the anti-MjPAN antibody based on Western analysis. The NTPase #1 formed a distinct complex based on native gel electrophoresis (Fig.15), and had a low but detectable ATPase activity of $1.58 \text{ nmol P}_i \text{ released} \cdot \text{min}^{-1} \cdot \text{mg protein}^{-1}$ (Table 10). The NTPase #1 hydrolyzed CTP at 69% the rate of ATP hydrolysis, and the ATPase activity was inhibited by equimolar concentrations of ADP to 43% of the original activity. The CTPase activity was only slightly inhibited by an equimolar concentration of CDP. Pre-incubation with NEM resulted in inhibition to 50% of the original activity. The NTPase #2 also formed a distinct complex based on native gel electrophoresis (Fig. 15) and hydrolyzed ATP at a rate of $3.14 \text{ nmol P}_i \text{ released} \cdot \text{min}^{-1} \cdot \text{mg protein}^{-1}$. Similar to NTPase #1, the ATPase activity of NTPase #2 was also inhibited by equimolar concentrations of ADP to 54% of the original activity (Table 11). The NTPase #2 also hydrolyzed CTP at 65% the rate of ATP hydrolysis, and this activity was not inhibited by an equimolar concentration of CDP. However, NTPase #2 was much more sensitive to inhibition by 5 mM NEM than NTPase #1, retaining only 7.5% of the activity measured in the absence of inhibitor (Table 11).

Purification and Characterization of a 20S Proteasome (Mm20S) from *M. maripaludis*

A 20S proteasome was identified (Fig. 14) and isolated from *M. maripaludis* cell free extract (Table 12). The molecular masses of the α and β subunits were estimated to be 28 and 24-kDa, respectively, based on SDS-PAGE (Fig. 16A), which is consistent

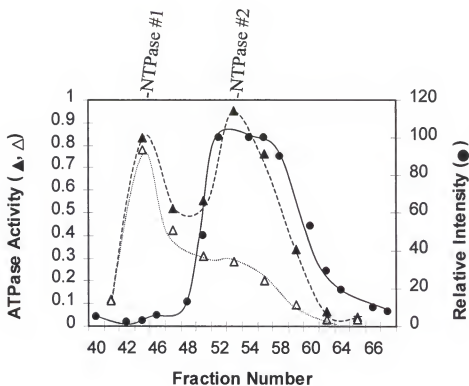


FIG. 14. Q Sepharose fractions during purification of a 20S proteasome and two NTPase proteins from *M.maripaludis* cell lysate. ATPase specific activity (nmol P_i per min mg^{-1} protein) in the absence (▲) and presence (△) of 10mM NEM.. (●) Relative Intensity measured by release of 7-amino-4-methylcoumarin from the fluorogenic substrate LLVY-AMC. NTPase peaks #1 and #2 are indicated. Note that NTPase #1 is not as inhibited by 10mM NEM as NTPase #2.

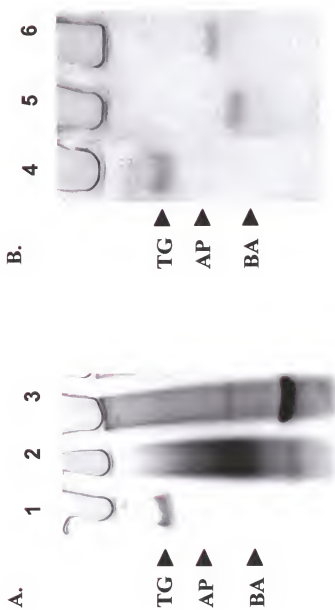


FIG. 15. Non-denaturing native gel electrophoresis of NTPase proteins from *M. maripaludis*. TG, thyroglobulin (669 kDa). AP, apoferritin (443 kDa). BA, β amylase (200 kDa). A. Gel stained with Coomassie blue. B. In-gel ATPase activity was measured at 37°C. Lanes 1 and 4, Purified PAN from *M. jannaschii*; Lanes 2 and 5, NTPase #1 from *M. maripaludis*; Lanes 3 and 6, NTPase #2 from *M. maripaludis*.

TABLE 10: Biochemical properties of NTPase #1 from *M. maripaludis*.

	Activity (nmol Pi min ⁻¹ mg ⁻¹)	% Control (no inhibitor)
<u>ATPase Activity</u>		
1 mM ATP only	1.58	100
1 mM ATP + 1 mM ADP	0.69	43
1 mM ATP + 5 mM NEM	0.81	50
<u>CTPase Activity</u>		
1 mM CTP only	1.09	100
1 mM CTP + 1 mM CDP	0.84	77

TABLE 11: Biochemical properties of NTPase #2 from *M. maripaludis*.

	Activity (nmol Pi min ⁻¹ mg ⁻¹)	% Control (no inhibitor)
<u>ATPase Activity</u>		
1 mM ATP only	3.14	100
1 mM ATP + 1 mM ADP	1.70	54
1 mM ATP + 5 mM NEM	0.24	7.5
<u>CTPase Activity</u>		
1 mM CTP only	2.07	100
1 mM CTP + 1 mM CDP	2.00	97

with the genome sequence of *M. maripaludis* LL. The 20S proteasome assembled into a 700-kDa complex based on Superose 6 gel filtration, and the subunits cross-reacted with anti-Mt α and anti-Mt β antibodies based on Western analyses (Fig. 16B). In addition, there were two bands detected on the Western blot probed with anti-Mt α antibody (Fig. 16B, Lane 3). There are two possible reasons for this. Either the β subunit is cross-reacting with the anti-Mt α antibody, or the α subunit is being cleaved to create a smaller protein. Pairwise sequence alignment of the β subunit of Mm20S to Mt α indicates that they are 32% identical (50% similar) at the amino acid level. Therefore, cross-reaction of the β subunit with the anti-Mt α antibody is possible. The purified 20S proteasome hydrolyzed the substrate LLVY-AMC at an optimum of 85°C (Fig. 17). In addition, the protein was stable when incubated at 80°C for 15 minutes before addition of the LLVY-AMC substrate, retaining 70% of the activity compared to a control sample incubated at 25°C (data not shown). While this temperature optimum, which is 43°C higher than the growth temperature optimum of the organism, was somewhat surprising, other 20S proteasomes have also demonstrated a higher optimum for peptide hydrolysis than the growth temperature of the organism (156, 286). The peptide-hydrolyzing activity was also measured using various synthetic substrates (Table 13). Of the substrates tested, the proteasome had the highest rate of cleavage carboxyl to the aromatic residue tyrosine using the substrate Suc-LLVY-AMC. The chymotrypsin-like (CL) activity was much lower when the substrates Suc-AAF-AMC or Suc-LY-AMC were used (Table 13). The proteasome also had significant PGPH activity using the substrate Cbz-LLE- β Na (Table 13). The activities observed against the substrates Suc-LLVY-AMC and Cbz-LLE- β Na were higher than any other previously characterized 20S proteasome except for the

TABLE 12. Purification of a 20S proteasome from *M. maripaludis*.

	Total Protein (mg)	Specific Activity	Total activity (nmol/min)	Purification (fold)	% Recovery
Cell free extract	518	0.70	363	1	100
Q Sepharose 26/10	64.4	0.74	47.6	1.1	13
PEG concentration	48.8	1.01	65.3	1.45	18
Superose 6	1.88	8.66	16.3	12.4	4.5
Hydroxyapatite	0.07	92.65	6.49	132.4	1.8
Mono Q HR 5/5	0.013	258	3.42	368	1.0

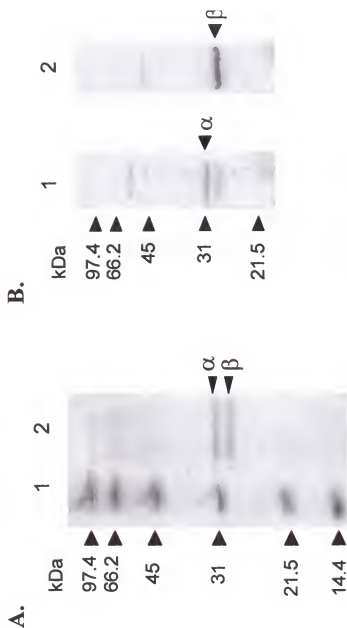


FIG. 16. SDS-PAGE gel and Western blot of the purified 20S proteasome from *M. maripaludis*. **A.** Purified 20S proteasome protein was separated by 12% reducing SDS-PAGE and stained with Coomassie blue. Lane 1: Molecular weight standards with masses indicated on left. Lane 2: 2 μ g purified 20S proteasome. The α and β subunits are indicated on the right. **B.** Proteins were separated by 12% reducing SDS-PAGE, electroblotted onto PVDF membrane, and analyzed using Western blotting techniques. Lane 1: Purified 20S proteasome probed with anti-Mta. Lane 2: Purified 20S proteasome probed with anti-Mif β . Molecular weight standards are indicated on the left.

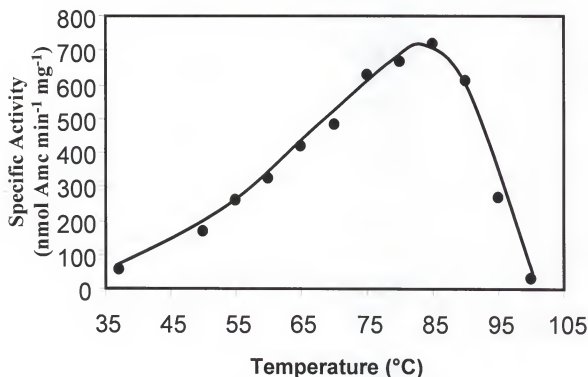


FIG. 17. Effect of temperature on the LLVY-AMC hydrolysis activity of the 20S proteasome from *M. maripaludis*. Purified 20S proteasome (0.1 μg per assay) was incubated in 25 mM Tris, 5 mM DTT at pH 7.5 using 20 μM LLVY-AMC as a substrate. Release of fluorogenic 7-amino-4-methylcoumarin was measured after incubation at the indicated temperatures. Data presented are the average of two experiments.

TABLE 13: Multiple peptide-hydrolyzing activities of the 20S proteasome from *M. maripaludis*.

Activity ^a	Substrate ^b	Specific Activity ^c
CL	Suc-LLVY-AMC	396.0
	Suc-AAF-AMC	87.4
	Suc-LY-AMC	14.9
PGPH	Cbz-LLE-βna	243.8
TL	Benz-FVR-AMC	31.3
	Boc-FSR-AMC	UD
SNAAP	Suc-IA-AMC	UD

^aCL, chymotrypsin-like; PGPH, peptidyl-glutamyl peptide hydrolyzing; TL, trypsin-like; SNAAP, small neutral amino acid preferring.

^bSuc, succinyl; AMC, 7-amido-4-methyl-coumarin; Cbz, carbobenzoxy; Benz, N-benzoyl; Boc, *t*-butyloxycarbonyl; βNa, β-naphthylamine.

^cSpecific activity was measured in micromoles of fluorogenic product · min⁻¹ · mg protein⁻¹. Peptide hydrolysis was measured at 60°C using 150 μM of the indicated substrates in 25 mM Tris buffer at pH 7.5 containing 5 mM DTT.

UD, undetectable.

recombinant *M. jannaschii* 20S proteasome (287). The trypsin-like (TL) activity of the Mm20S proteasome using the substrate Benz-FVR-AMC was low but virtually identical to the activity previously observed for the Mj20S proteasome (287). The Mm20S proteasome had no detectable activity when the substrate Boc-FSR-AMC was used. This was not surprising, as low TL activities have also been observed for other archaeal and eubacterial proteasomes (6, 156, 299), in contrast to mammalian 20S proteasomes which generally have high TL activity (reviewed in 197). Neither the Mm20S proteasome (Table 13) or the Mj20S proteasome (287) had any detectable small neutral amino acid preferring (SNAAP) activity. The LLVY-AMC hydrolyzing activity of the Mm20S proteasome was also slightly stimulated by NaCl, KCl, and MgCl₂ salts by 1.6-, 1.4-, and 1.3-fold, respectively (data not shown). Electron microscopy of the 700-kDa Mm20S proteasome reveals a four-ringed structure of 12 by 17 nm with a central channel (Fig. 18). This is consistent with other archaeal 20S proteasomes which have been purified including those from *M. thermophila* (156) and *T. acidophilum* (41).

Synthesis of MmPAN and MjPAN in Recombinant *M. maripaludis*

The genes expressing His₆-tagged MjPAN and MmPAN as well as untagged constructs were placed into the expression shuttle vector pWLG40 and the proteins were synthesized in recombinant *M. maripaludis* (See Table 4 for plasmid descriptions). The strains producing MjPAN or MmPAN were grown to mid-log phase and harvested. The MjPAN and MmPAN proteins that were synthesized in the recombinant strains were detected by Western blot of the cell lysate (Fig. 19). When either MmPAN or MjPAN were synthesized as untagged constructs or with C-terminal His₆ tags, overproduction was easily detected by Western blot of the cell lysate (Fig. 19, Lanes 3,5,6,7, and 9). In



FIG. 18. Transmission electron microscopy of purified 20S proteasomes from *M. maripaludis*. Side views in which the four-ringed structure is visible, and end-on views (inset) can be seen.

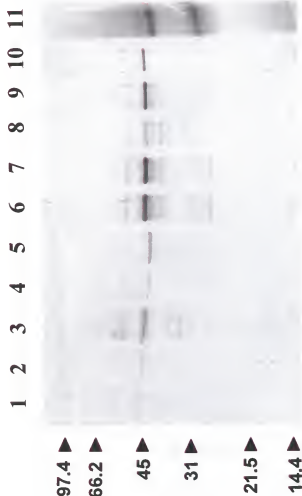


FIG. 19. Recombinant *M. maripaludis* strains producing MmPAN. Proteins were separated by 12% reducing SDS-PAGE and analyzed by Western blot using anti-MjPAN antibody. Lane 1: 10 μ g *M. maripaludis* JJ cell lysate (no plasmid). Lanes 2-5: Cell lysate (10 μ g) of *M. maripaludis* (pWLG40+*lacZ*), (pJAM738), (pJAM739), and (pJAM737), respectively. Lane 6-9: Cell lysate (1 μ g) of *M. maripaludis* (pJAM740), (pJAM741), (pJAM743), and (pJAM742), respectively. Lane 10: 0.05 μ g purified MjPAN. Lane 11: Cell lysate of recombinant *E. coli* (pMjPAN2) cells producing MjPAN proteins of 47-, 39-, and 30-kDa. Molecular weight standards are indicated on the left.

contrast, there was no obvious overproduction of MmPAN when synthesized as an N-terminal His₆-tagged construct, and only slight overproduction of N- His₆-tagged MjPAN could be detected (Fig. 19, Lanes 4 and 8). Interestingly, there was no detectable synthesis of truncated MjPAN(Δ 1-73) even when the ORF on the plasmid contained the wild-type internal translation start site (Fig. 19, Lane 6). This indicates that *M. maripaludis* cell proteins are not initiating synthesis from the internal translation start site when grown under these conditions.

Cell lysate of the *M. maripaludis* strains producing PAN proteins was analyzed for ATPase activity (Table 14). In general, the strains producing MmPAN had little if any increase in ATPase activity compared to the wild-type strain. However, the cell lysate from the *M. maripaludis* strains producing MjPAN had a significant increase of ATPase activity compared to the wild-type strain (Table 14). When assayed at 50°C, the ATPase activity of wild-type *M. maripaludis*, strains producing MmPAN, and strains producing MjPAN with N-terminal His₆-tags showed little change in ATPase activity. In contrast, the ATPase activity of the cell lysate of *M. maripaludis* producing untagged MjPAN increased to 4-fold the activity of the wild-type strain when incubated at 50°C (Table 14). These results are consistent with recombinant *E. coli* studies which also showed that addition of an N-terminal His₆-tag can reduce the thermostability of MjPAN.

Purification of His₆-tagged MmPAN from Recombinant *M. maripaludis*

The C-terminally His₆-tagged MmPAN from recombinant *M. maripaludis* (pJAM738) bound to a Cu²⁺-Sepharose affinity column and was eluted using 500 mM imidazole buffer (Fig. 20). When the MmPAN protein fraction was applied to a Superose 6 column, the protein eluted in a very broad peak, and ATP hydrolysis could not be

TABLE 14. Comparison of ATPase activities of cell lysate of recombinant *M. maripaludis* strains.

Plasmid	Protein Overproduced	% Activity ^a	
		37°C	50°C
None	None	100 ^b	100 ^c
pWLG40 + <i>lacZ</i>	β -galactosidase	103	103
pJAM738	C-His ₆ -MmPAN	103	94
pJAM739	N-His ₆ -MmPAN	119	146
pJAM737	MmPAN	126	145
pJAM741	C-His ₆ -MjPAN	124	183
pJAM742	C-His ₆ -MjPAN	240	328
pJAM743	N-His ₆ -MjPAN	130	114
pJAM740	MjPAN	162	416

^a%Activity is the ATPase activity compared to cell lysate from wild-type *M. maripaludis* without any plasmid. Cell lysate was assayed in 20 mM Tris buffer at pH 7.5 containing 100 mM NaCl, 10 mM MgCl₂, and 1 mM ATP.

^b100% Activity, 9.7 nmol P_i released · min⁻¹ · mg protein⁻¹ at 37°C.

^c100% Activity, 8.5 nmol P_i released · min⁻¹ · mg protein⁻¹ at 50°C.

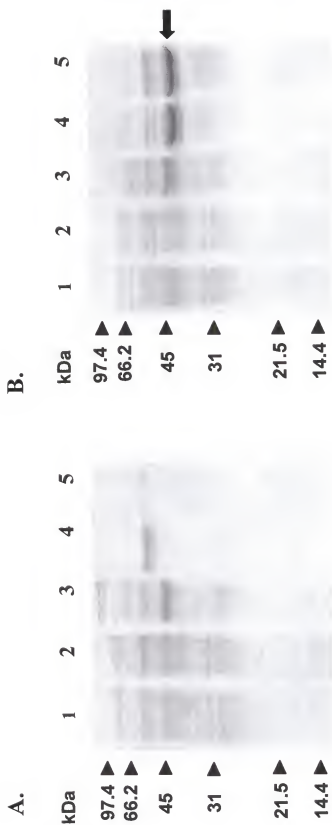


FIG. 20. Purification of MmPAN from recombinant *M. maripaludis* using a Cu^{2+} -Sephacrose column. Proteins from Cu^{2+} -Sephacrose column fractions were separated by 12% reducing SDS-PAGE and stained with Coomassie blue. **A.** *M. maripaludis* (pWLG40 + *lacZ*) **B.** *M. maripaludis* (pJAM738). Lane 1: *M. maripaludis* cell lysate; Lane 2: 5 mM imidazole buffer wash; Lane 3: 60 mM imidazole buffer wash; Lanes 4 and 5: 500 mM imidazole buffer elution fractions. Imidazole buffers contained 20 mM Tris, pH 7.5, 100 mM NaCl, 10% glycerol, and indicated concentration of imidazole. Arrow on right indicates MmPAN. Molecular masses are indicated on the left.

detected. These results were similar to those obtained during purification of the C-terminal His₆-tagged MmPAN from recombinant *E. coli*. Taken together, these results suggest that the C-terminal region may be important in formation of the MmPAN complex, or that the C-terminal His₆-tag may disrupt complex formation.

SUMMARY AND CONCLUSIONS

The results of this study provide evidence that PAN and 20S proteasome proteins are produced in both of the methanogenic archaea *M. jannaschii* and *M. maripaludis*. This study also examined the biochemical and physical properties of the *M. jannaschii* PAN protein. Structural characterization of MjPAN by electron microscopy, gel filtration, and non-denaturing native gel electrophoresis showed that the protein forms a 550 kDa complex of approximately 12 subunits which is stable in the absence of ATP. In addition, partial purification of MjPAN from *M. jannaschii* reveals that it also associates into a 550 kDa complex with similar biochemical properties as the recombinant protein. Separate biochemical characterization of the full-length MjPAN protein and the N-terminally truncated MjPAN(Δ 1-73) protein revealed that the highly charged N-terminus of MjPAN modulates the rate of ATP and CTP hydrolysis and provides nucleotide specificity to the active site. In addition, the N-terminus appeared to influence the sensitivity of MjPAN to inhibitors such as NEM and PCMS as well as affect product inhibition by ADP or CDP. Secondary structure analysis suggests that an α -helical, putative coiled-coil region in the N-terminus of MjPAN is absent in the truncated protein and, therefore, may play an important role in stabilizing MjPAN at extreme temperatures. Similar to other proteins which are stable at extreme temperature and/or salt conditions, it is possible that ion-pairs, hydrophobic clusters, or other structural elements are required for the extreme thermostability of MjPAN and that these are disrupted by removal of the

N-terminal peptide. Taken together, the results of this study suggest that the full-length MjPAN protein, detected in *M. jannaschii* using Western blot analysis, is likely the native form which functions at the extreme growth temperature of this organism.

This study also demonstrated that MjPAN stimulates the proteolytic activity of archaeal 20S proteasomes in an ATP- or CTP-dependent manner. In addition, electron micrographs suggested that MjPAN and the 20S proteasome associate *in vitro* as a complex of the 20S proteolytic core capped either singly or doubly with the MjPAN protein complex. This complex resembled the proteasome Base subcomplex recently described for yeast (72).

The mesophilic *M. maripaludis* is a methanoarchaeon for which genetic methods and high-level culture techniques have been developed. Therefore, this organism was chosen as an attractive model in which to continue study of ATP-dependent proteolysis in the Archaea. A 20S proteasome was purified from *M. maripaludis* which had a surprisingly high temperature optimum of 80°C for hydrolysis of peptide substrates, compared to its growth temperature optimum of 37°C. Two NTPase complexes were also partially purified from *M. maripaludis*. Both had NEM-inhibitable ATPase and CTPase activity, characteristic of AAA-ATPases including MjPAN. However, neither NTPase cross-reacted with the MjPAN antibody. Therefore, a *M. maripaludis* genomic library was constructed, and an ORF with 70% identity to the *M.jannaschii pan* gene was isolated. The Mmp_{pan} ORF was cloned into expression vectors, and MmPAN was purified from recombinant *E. coli*. The purified MmPAN protein formed a dodecameric complex, but had no ATPase or CTPase activity and was unstable at 50°C. This suggested that the recombinant MmPAN may not be folding properly in the cytosol of

recombinant *E. coli*. Thus, to further characterize PAN, both MmPAN and MjPAN were over-produced in recombinant *M. maripaludis* as C- and N-terminally His₆-tagged proteins as well as untagged proteins. The purified C-terminally His₆-tagged MmPAN had biochemical properties similar to the MmPAN protein produced in recombinant *E. coli*. Sequence alignment of MjPAN and MmPAN revealed that MmPAN lacked the highly charged amino-terminus of MjPAN. This may explain the lack of thermostability of recombinant MmPAN. The reason MmPAN lacks ATPase activity remains to be determined.

Current and future studies of this project are also focused on determining proteins which may interact with PAN and the 20S proteasome *in vivo*. Using the His₆-tagged constructs of PAN from this study may identify proteins which interact with the proteasome *in vivo*. This is based on a recent study of the yeast proteasome which used affinity-tagged constructs to rapidly purify intact 26S proteasomes and RP (19S cap) complexes from the cytosol (267). Mass spectrometry techniques were then used to identify 24 previously undiscovered proteins which interacted with the 26S proteasome.

The recently completed genome sequence of *M. maripaludis*, combined with the genetic techniques which have been developed for use in this methanogenic microorganism, should also aid future studies of the proteasome in Archaea. Further characterization of MjPAN and MmPAN produced in recombinant *M. maripaludis* should reveal more about the similarities and differences between these two proteins, one from a mesophile and the other from a hyper-thermophile.

REFERENCES

- 1 Akiyama, K., S. Kagawa, T. Tamura, N. Shimbara, M. Takashina, P. Kristensen, K. B. Hendil, K. Tanaka, and A. Ichihara. 1994. Replacement of proteasome subunits X and Y by LMP7 and LMP2 induced by interferon-gamma for acquirement of the functional diversity responsible for antigen processing. *FEBS Lett.* **343**:85-88.
- 2 Akiyama, Y., and K. Ito. 2000. Roles of multimerization and membrane association in the proteolytic functions of FtsH (HflB). *EMBO J.* **19**:3888-3895.
- 3 Akiyama, Y., A. Kihara, and K. Ito. 1996. Subunit *a* of proton ATPase F_0 sector is a substrate of the FtsH protease in *Escherichia coli*. *FEBS Lett.* **399**:26-28.
- 4 Akiyama, Y., A. Kihara, H. Tokuda, and K. Ito. 1996. FtsH (HflB) is an ATP-dependent protease selectively acting on SecY and some other membrane proteins. *J.Biol.Chem.* **271**:31196-31201.
- 5 Akiyama, Y., T. Yoshihisa, and K. Ito. 1995. FtsH, a membrane-bound ATPase, forms a complex in the cytoplasmic membrane of *Escherichia coli*. *J.Biol.Chem.* **270**:23485-23490.
- 6 Akopian, T. N., A. F. Kisselev, and A. L. Goldberg. 1997. Processive degradation of proteins and other catalytic properties of the proteasome from *Thermoplasma acidophilum*. *J.Biol.Chem.* **272**:1791-1798.
- 7 Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J.Mol.Biol.* **215**:403-410.
- 8 Anfinsen, C. B. 1967. The formation of the tertiary structure of proteins. *Harvey Lect.* **61**:95-116.
- 9 Arendt, C. S., and M. Hochstrasser. 1999. Eukaryotic 20S proteasome catalytic subunit propeptides prevent active site inactivation by N-terminal acetylation and promote particle assembly. *EMBO J.* **18**:3575-3585.
- 10 Arrigo, A.-P., K. Tanaka, A. L. Goldberg, and W. J. Welch. 1988. Identity of the 19S 'prosome' particle with the large multifunctional protease complex of mammalian cells (the proteasome). *Nature* **331**:192-194.
- 11 Babst, M., T. K. Sato, L. M. Banta, and S. D. Emr. 1997. Endosomal transport function in yeast requires a novel AAA-type ATPase, Vps4p. *EMBO J.* **16**:1820-1831.
- 12 Balow R.M., Ragnarsson U., and Zetterqvist O. 1983. Tripeptidyl aminopeptidase in the extralysosomal fraction of rat liver. *J.Biol.Chem.* **258**:11622-11628.

- 13 **Bauer, M. W., S. B. Halio, and R. M. Kelly.** 1997. Purification and characterization of a proteasome from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Appl. Environ. Microbiol.* **63**:1160-1164.
- 14 **Beneke, S., H. Bestgen, and A. Klein.** 1995. Use of the *Escherichia coli uidA* gene as a reporter in *Methanococcus voltae* for the analysis of the regulatory function of the intergenic region between the operons encoding selenium-free hydrogenases. *Mol. Gen. Genet.* **248**:225-228.
- 15 **Berghofer Y., Agha-Amiri K., and Klein A.** 1994. Selenium is involved in the negative regulation of the expression of selenium-free [NiFe] hydrogenases in *Methanococcus voltae*. *Mol. Gen. Genet.* **242**:369-373.
- 16 **Berghofer, Y., and A. Klein.** 1995. Insertional mutations in the hydrogenase *vhc* and *frc* operons encoding selenium-free hydrogenases in *Methanococcus voltae*. *Appl. Environ. Microbiol.* **61**:1770-1775.
- 17 **Bertani, G., and L. Baresi.** 1987. Genetic transformation in the methanogen *Methanococcus voltae* PS. *J. Bacteriol.* **169**:2730-2738.
- 18 **Beuron, F., M. R. Maurizi, D. M. Belnap, E. Kocsis, F. P. Booy, M. Kessel, and A. C. Steven.** 1998. At sixes and sevens: characterization of the symmetry mismatch of the ClpAP chaperone-assisted protease. *J. Struct. Biol.* **123**:248-259.
- 19 **Bochtler, M., L. Ditzel, M. Groll, and R. Huber.** 1997. Crystal structure of heat shock locus V (HslV) from *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **94**:6070-6074.
- 20 **Boone, D. R., R. L. Johnson, and Y. Liu.** 1989. Diffusion of the interspecies electron carriers H_2 and formate in methanogenic ecosystems and its implications in the measurement of K_m for H_2 or formate uptake. *Appl. Environ. Microbiol.* **55**:1735-1741.
- 21 **Borges, A. C., and S. L. Gomes.** 2000. PEST sequences in cAMP-dependent protein kinase subunits of the aquatic fungus *Blastocladiella emersonii* are necessary for *in vitro* degradation by endogenous proteases. *Mol. Microbiol.* **36**:926-939.
- 22 **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**:248-254.
- 23 **Brannigan, J. A., G. Dodson, H. J. Duggleby, P. C. E. Moody, J. L. Smith, D. R. Tomchick, and A. G. Murzin.** 1995. A protein catalytic framework with an N-terminal nucleophile is capable of self-activation. *Nature* **378**:416-419.

- 24 **Braun, B. C., M. Glickman, R. Kraft, B. Dahlmann, P. M. Kloetzel, D. Finley, and M. Schmidt.** 1999. The base of the proteasome regulatory particle exhibits chaperone-like activity. *Nat.Cell Biol.* 1:221-226.
- 25 **Brooks, P., G. Fuertes, R. Z. Murray, S. Bose, E. Knecht, M. C. Rechsteiner, K. B. Hendil, K. Tanaka, J. Dyson, and J. Rivett.** 2000. Subcellular localization of proteasomes and their regulatory complexes in mammalian cells. *Biochem J.* 346 Pt 1:155-161.
- 26 **Bult, C. J., O. White, G. J. Olsen, L. Zhou, D. Fleischmann, G. G. Sutton, J. A. Blake, M. FitzGerald, R. A. Clayton, J. D. Gocayne, A. R. Kerlavage, B. A. Dougherty, J.-F. Tomb, M. D. Adams, C. I. Reich, R. Overbeek, E. F. Kirkness, K. G. Weinstock, J. M. Merrick, A. Glodek, J. L. Scott, N. S. M. Geoghagen, J. F. Weidman, J. L. Fuhrmann, D. Nguyen, T. R. Utterback, J. M. Kelley, J. D. Peterson, P. W. Sadow, M. C. Hanna, M. D. Cotton, K. M. Roberts, M. A. Hurst, B. P. Kaine, M. Borodovsky, H.-P. Klenk, C. M. Fraser, H. O. Smith, C. R. Woese, and J. C. Venter.** 1996. Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science* 273:1058-1073.
- 27 **Burke, M., K. N. Rajasekharan, S. Maruta, and M. Ikebe.** 1990. A second consensus sequence of ATP-requiring proteins resides in the 21- kDa C-terminal segment of myosin subfragment 1. *FEBS Lett.* 262:185-188.
- 28 **Burri, L., J. Hockendorff, U. Boehm, T. Klamp, R. J. Dohmen, and F. Levy.** 2000. Identification and characterization of a mammalian protein interacting with 20S proteasome precursors. *Proc.Natl.Acad.Sci.U.S.A.* 97:10348-10353.
- 29 **Cann, I. K. O., and Y. Ishino.** 1999. Archaeal DNA replication: identifying the pieces to solve a puzzle. *Genetics* 152:1249-1267.
- 30 **Chen, P., and M. Hochstrasser.** 1995. Biogenesis, structure, and function of the yeast 20S proteasome. *EMBO J.* 14:2620-2630.
- 31 **Chen, P., and M. Hochstrasser.** 1996. Autocatalytic subunit processing couples active site formation in the 20S proteasome to completion of assembly. *Cell* 86:961-972.
- 32 **Chen, Z., J. Hagler, V. J. Palombella, F. Melandri, D. Scherer, D. Ballard, and T. Maniatis.** 1995. Signal-induced site-specific phosphorylation targets I κ B α ubiquitin-proteasome pathway. *Genes Dev.* 9:1586-1597.
- 33 **Chen, Z. J., L. Parent, and T. Maniatis.** 1996. Site-specific phosphorylation of I κ B α by a novel ubiquitination-dependent protein kinase activity. *Cell* 84:853-862.
- 34 **Chernoff, Y. O., S. L. Lindquist, B. Ono, S. G. Inge-Vechtomov, and S. W. Liebman.** 1995. Role of the chaperone protein Hsp104 in propagation of the yeast prion- like factor [psi⁺]. *Science* 268:880-884.

- 35 **Chin, D. T., S. A. Goff, T. Webster, T. Smith, and A. L. Goldberg.** 1988. Sequence of the *lon* gene in *Escherichia coli*. A heat-shock gene which encodes the ATP-dependent protease La. *J.Biol.Chem.* **263**:11718-11728.
- 36 **Ciechanover, A.** 1998. The ubiquitin-proteasome pathway: on protein death and cell life. *EMBO J.* **17**:7151-7160.
- 37 **Confalonieri, F., and M. Duguët.** 1995. A 200-amino acid ATPase module in search of a basic function. *BioEssays* **17**:639-650.
- 38 **Confalonieri, F., J. Marsault, and M. Duguët.** 1994. SAV, an archaeobacterial gene with extensive homology to a family of highly conserved eukaryotic ATPases. *J.Mol.Biol.* **235**:396-401.
- 39 **Cornish-Bowden, A.** 1995. *Fundamentals of Enzyme Kinetics*. Portland Press, London.
- 40 **Coux, O., K. Tanaka, and A. L. Goldberg.** 1996. Structure and functions of the 20S and 26S proteasomes. *Annu.Rev.Biochem.* **65**:801-847.
- 41 **Dahlmann, B., F. Kopp, L. Kuehn, B. Nidel, G. Pfeifer, R. Hegerl, and W. Baumeister.** 1989. The multicatalytic proteinase (prosome) is ubiquitous from eukaryotes to archaeobacteria. *FEBS Lett.* **251**:125-131.
- 42 **DeMartino, G. N., C. R. Moomaw, O. P. Zagnitko, R. J. Proske, M. Chu-Ping, S. J. Afendis, J. C. Swaffield, and C. A. Slaughter.** 1994. PA700, an ATP-dependent activator of the 20 S proteasome, is an ATPase containing multiple members of a nucleotide-binding protein family. *J.Biol.Chem.* **269**:20878-20884.
- 43 **DeMartino, G. N., R. J. Proske, C. R. Moomaw, A. A. Strong, X. Song, H. Hisamatsu, K. Tanaka, and C. A. Slaughter.** 1996. Identification, purification, and characterization of a PA700-dependent activator of the proteasome. *J.Biol.Chem.* **271**:3112-3118.
- 44 **Deveraux, Q., V. Ustrell, C. Pickart, and M. Rechsteiner.** 1994. A 26S protease subunit that binds ubiquitin conjugates. *J.Biol.Chem.* **269**:7059-7061.
- 45 **Ditzel, L., J. Löwe, D. Stock, K. O. Stetter, H. Huber, R. Huber, and S. Steinbacher.** 1998. Crystal structure of the thermosome, the archaeal chaperonin and homolog of CCT. *Cell* **93**:125-138.
- 46 **Driscoll, J., M. G. Brown, D. Finley, and J. J. Monaco.** 1993. MHC-linked *LMP* gene products specifically alter peptidase activities of the proteasome. *Nature* **365**:262-264.
- 47 **Dubiel, W., K. Ferrell, G. Pratt, and M. Rechsteiner.** 1992. Subunit 4 of the 26S protease is a member of a novel eukaryotic ATPase family. *J.Biol.Chem.* **267**:22699-702.

- 48 Dubiel, W., K. Ferrell, and M. Rechsteiner. 1995. Subunits of the regulatory complex of the 26S protease. *Mol.Biol.Rep.* **21**:27-34.
- 49 Duggleby, H. J., S. P. Tolley, C. P. Hill, E. J. Dodson, G. Dodson, and P. C. Moody. 1995. Penicillin acylase has a single-amino-acid catalytic centre. *Nature* **373**:264-268.
- 50 Edgell D.R., and Doolittle W.F. 1997. Archaea and the origin(s) of DNA replication proteins. *Cell* **89**:995-998.
- 51 Edman, P. 1970. Sequence determination. *Mol.Biol.Biochem.Biophys.* **8**:211-255:211-255.
- 52 Edmunds, T., and A. L. Goldberg. 1986. Role of ATP hydrolysis in the degradation of proteins by protease La from *Escherichia coli*. *J.Cell Biochem.* **32**:187-191.
- 53 Ellis, R. J., and F. U. Hartl. 1996. Protein folding in the cell: competing models of chaperonin function. *FASEB J.* **10**:20-26.
- 54 Feiler, H. S., T. Desprez, V. Santoni, J. Kronenberger, M. Caboche, and J. Traas. 1995. The higher plant *Arabidopsis thaliana* encodes a functional CDC48 homologue which is highly expressed in dividing and expanding cells. *EMBO J.* **14**:5626-5637.
- 55 Finley, D., K. Tanaka, C. Mann, H. Feldmann, M. Hochstrasser, R. Vierstra, S. Johnston, R. Hampton, J. Haber, J. Mccusker, P. Silver, L. Frontali, P. Thorsness, A. Varshavsky, B. Byers, K. Madura, S. I. Reed, D. Wolf, S. Jentsch, T. Sommer, W. Baumeister, A. Goldberg, V. Fried, D. M. Rubin, and A. Toh-e. 1998. Unified nomenclature for subunits of the *Saccharomyces cerevisiae* proteasome regulatory particle. *Trends Biochem.Sci.* **23**:244-245.
- 56 Foss, G. S., F. Larsen, J. Solheim, and H. Prydz. 1998. Constitutive and interferon- γ -induced expression of the human proteasome subunit multicatalytic endopeptidase complex-like 1. *Biochim.Biophys.Acta* **1402**:17-28.
- 57 Frohlich, K. U., H. W. Fries, J. M. Peters, and D. Mecke. 1995. The ATPase activity of purified CDC48p from *Saccharomyces cerevisiae* shows complex dependence on ATP-, ADP-, and NADH- concentrations and is completely inhibited by NEM. *Biochim.Biophys.Acta* **1253**:25-32.
- 58 Fröh, K., M. Gossen, K. Wang, H. Bujard, P. A. Peterson, and Y. Yang. 1994. Displacement of housekeeping proteasome subunits by MHC-encoded LMPs: a newly discovered mechanism for modulating the multicatalytic proteinase complex. *EMBO J.* **13**:3236-3244.

- 59 Gaczynska, M., K. L. Rock, and A. L. Goldberg. 1993. γ -interferon and expression of MHC genes regulate peptide hydrolysis by proteasomes. *Nature* **365**:264-267.
- 60 Ganoth, D., E. Leshinsky, E. Eytan, and A. Hershko. 1988. A multicomponent system that degrades proteins conjugated to ubiquitin. Resolution of factors and evidence for ATP-dependent complex formation. *J.Biol.Chem.* **263**:12412-12419.
- 61 Gao, Y., J. O. Thomas, R. L. Chow, G. H. Lee, and N. J. Cowan. 1992. A cytoplasmic chaperonin that catalyzes beta-actin folding. *Cell* **69**:1043-1050.
- 62 Gardner, W. L. 2000. Expression vectors for the methane-producing Archaeon *Methanococcus maripaludis*. Ph.D. Dissertation, University of Georgia.
- 63 Gardner, W. L., and W. B. Whitman. 1999. Expression vectors for *Methanococcus maripaludis*. Overexpression of acetohydroxyacid synthase and beta-galactosidase. *Genetics* **152**:1439-1447.
- 64 Geier, E., G. Pfeifer, M. Wilm, M. Lucchiari-Hartz, W. Baumeister, K. Eichmann, and G. Niedermann. 1999. A giant protease with potential to substitute for some functions of the proteasome. *Science* **283**:978-981.
- 65 Geissler, S., K. Siegers, and E. Schiebel. 1998. A novel protein complex promoting formation of functional alpha- and gamma-tubulin. *EMBO J.* **17**:952-966.
- 66 Gerards, W. L., W. W. de Jong, H. Bloemendal, and W. Boelens. 1998. The human proteasomal subunit HsC8 induces ring formation of other α -type subunits. *J.Mol.Biol.* **275**:113-121.
- 67 Gerards, W. L., J. Enzlin, M. Haner, I. L. Hendriks, U. Aeby, H. Bloemendal, and W. Boelens. 1997. The human α -type proteasomal subunit HsC8 forms a double ringlike structure, but does not assemble into proteasome-like particles with the β -type subunits HsDelta or HsBPROS26. *J.Biol.Chem.* **272**:10080-10086.
- 68 Gernhardt, P., O. Possot, M. Foglino, L. Sibold, and A. Klein. 1990. Construction of an integration vector for use in the archaeobacterium *Methanococcus voltae* and expression of a eubacterial resistance gene. *Mol.Gen.Genet.* **221**:273-279.
- 69 Ghoda, L., D. Sidney, M. Macrae, and P. Coffino. 1992. Structural elements of ornithine decarboxylase required for intracellular degradation and polyamine-dependent regulation. *Mol.Cell Biol.* **12**:2178-2185.
- 70 Ghoda, L., W. T. van Daalen, M. Macrae, D. Ascherman, and P. Coffino. 1989. Prevention of rapid intracellular degradation of ODC by a carboxyl-terminal truncation. *Science* **243**:1493-1495.

- 71 Glas, R., M. Bogoy, J. S. McMaster, M. Gaczynska, and H. L. Ploegh. 1998. A proteolytic system that compensates for loss of proteasome function. *Nature* **392**:618-622.
- 72 Glickman, M. H., D. M. Rubin, O. Coux, I. Wefes, G. Pfeifer, Z. Cjeka, W. Baumeister, V. A. Fried, and D. Finley. 1998. A subcomplex of the proteasome regulatory particle required for ubiquitin-conjugate degradation and related to the COP9-signalosome and eIF3. *Cell* **94**:615-623.
- 73 Glickman, M. H., D. M. Rubin, V. A. Fried, and D. Finley. 1998. The regulatory particle of the *Saccharomyces cerevisiae* proteasome. *Mol. Cell Biol.* **18**:3149-3162.
- 74 Glickman, M. H., D. M. Rubin, H. Fu, C. N. Larsen, O. Coux, I. Wefes, G. Pfeifer, Z. Cjeka, R. Vierstra, W. Baumeister, V. Fried, and D. Finley. 1999. Functional analysis of the proteasome regulatory particle. *Mol. Biol. Rep.* **26**:21-28.
- 75 Gohl, H. P., B. Grondahl, and M. Thomm. 1995. Promoter recognition in archaea is mediated by transcription factors: identification of transcription factor aTFB from *Methanococcus thermolithotrophicus* as archaeal TATA-binding protein. *Nucleic Acids Res.* **23**:3837-3841.
- 76 Golbik, R., A. N. Lupas, K. K. Koretke, W. Baumeister, and J. Peters. 1999. The Janus face of the archaeal Cdc48/p97 homologue VAT: protein folding versus unfolding. *Biol. Chem.* **380**:1049-1062.
- 77 Gonciarz-Swiatek, M., A. Wawrzynow, S. J. Um, B. A. Learn, R. McMacken, W. L. Kelley, C. Georgopoulos, O. Sliemers, and M. Zylicz. 1999. Recognition, targeting, and hydrolysis of the λ O replication protein by the ClpP/ClpX protease. *J. Biol. Chem.* **274**:13999-14005.
- 78 Gottesman, S. 1999. Regulation by proteolysis: developmental switches. *Curr. Opin. Microbiol.* **2**:142-147.
- 79 Gottesman, S., E. Roche, Y. Zhou, and R. T. Sauer. 1998. The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system. *Genes Dev.* **12**:1338-1347.
- 80 Gottesman, S., S. Wickner, and M. R. Maurizi. 1997. Protein quality control: triage by chaperones and proteases. *Genes Dev.* **11**:815-823.
- 81 Gregory, D. W., and B. J. Pirie. 1973. Wetting agents for biological electron microscopy: I. General considerations and negative staining. *J. Microsc.* **99**:251-255.
- 82 Grimaud, R., M. Kessel, F. Beuron, A. C. Steven, and M. R. Maurizi. 1998. Enzymatic and structural similarities between the *Escherichia coli* ATP-dependent proteases, ClpXP and ClpAP. *J. Biol. Chem.* **273**:12476-12481.

- 83 Groll, M., M. Bajorek, A. Kohler, L. Moroder, D. M. Rubin, R. Huber, M. H. Glickman, and D. Finley. 2000. A gated channel into the proteasome core particle. *Nat.Struct.Biol.* 7:1062-1067.
- 84 Groll, M., L. Ditzel, J. Löwe, D. Stock, M. Bochtler, H. D. Bartunik, and R. Huber. 1997. Structure of 20S proteasome from yeast at 2.4 Å resolution. *Nature* 386 :463-471.
- 85 Gross-Mesilaty, S., J. L. Hargrove, and A. Ciechanover. 1997. Degradation of tyrosine aminotransferase (TAT) via the ubiquitin- proteasome pathway. *FEBS Lett.* 405:175-180.
- 86 Gutsche, I., L. O. Essen, and W. Baumeister. 1999. Group II chaperonins: new TRiC(k)s and turns of a protein folding machine. *J.Mol.Biol.* 293:295-312.
- 87 Halboth S., and Klein A. 1992. *Methanococcus voltae* harbors four gene clusters potentially encoding two [NiFe] and two [NiFeSe] hydrogenases, each of the cofactor F₄₂₀-reducing or F₄₂₀-non-reducing types. *Mol.Gen.Genet.* 233:217-224.
- 88 Halio, S. B., R. M. Kelly, I. I. Blumentals, S. A. Short, and B. M. Merrill. 1996. Sequence, expression in *Escherichia coli*, and analysis of the gene encoding a novel intracellular protease (PfpI) from the hyperthermophilic archaeon *Pyrococcus furiosus*. *J.Bacteriol.* 178:2605-2612.
- 89 Hanson, P. I., R. Roth, H. Morisaki, R. Jahn, and J. E. Heuser. 1997. Structure and conformational changes in NSF and its membrane receptor complexes visualized by quick-freeze/deep-etch electron microscopy. *Cell* 90:523-535.
- 90 Harlow, E., and D. Lane. 1999. Using Antibodies: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- 91 Harris, J. R. 1968. Release of a macromolecular protein component from human erythrocyte ghosts. *Biochim.Biophys.Acta* 150:534-537.
- 92 Heilig, R. 2000. *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome structure and evolution. Genbank accession **AJ248283 to AJ248288**.
- 93 Herman, C., D. Thevenet, P. Boulloc, G. C. Walker, and R. D'Ari. 1998. Degradation of carboxy-terminal-tagged cytoplasmic proteins by the *Escherichia coli* protease HflB (FtsH). *Genes Dev.* 12:1348-1355.
- 94 Hershko, A., and A. Ciechanover. 1998. The ubiquitin system. *Annu.Rev.Biochem.* 67:425-480.
- 95 Hochstrasser, M. 1996. Ubiquitin-dependent protein degradation. *Annu.Rev.Genet.* 30:405-439.

- 96 Hochstrasser, M., F. R. Papa, P. Chen, S. Swaminathan, P. Johnson, L. Stillman, A. Y. Amerik, and S.-J. Li. 1995. The DOA pathway: studies on the functions and mechanisms of ubiquitin-dependent protein degradation in the yeast *Saccharomyces cerevisiae*. Cold Spring Harb. Symp. Quant. Biol. LX:503-513.
- 97 Hoffman, L., and M. Rechsteiner. 1994. Activation of the multicatalytic protease. The 11S regulator and 20S ATPase complexes contain distinct 30-kilodalton subunits. J. Biol. Chem. 269:16890-16895.
- 98 Hoffman, L., and M. Rechsteiner. 1996. Nucleotidase activities of the 26S proteasome and its regulatory complex. J. Biol. Chem. 271:32538-32545.
- 99 Hofman-Bang, J., M. Lange, E. Conway de Macario, A. J. Macario, and B. K. Ahring. 1999. The genes coding for the hsp70 (*dnaK*) molecular chaperone machine occur in the moderate thermophilic archaeon *Methanosarcina thermophila* TM-1. Gene 238:387-395.
- 100 Hoskins, J. R., M. Pak, M. R. Maurizi, and S. Wickner. 1998. The role of the ClpA chaperone in proteolysis by ClpAP. Proc. Natl. Acad. Sci. U.S.A. 95:12135-12140.
- 101 Hough, R., G. Pratt, and M. Rechsteiner. 1986. Ubiquitin-lysosome conjugates. Identification and characterization of an ATP-dependent protease from rabbit reticulocyte lysates. J. Biol. Chem. 261:2400-2408.
- 102 Hough, R., and M. Rechsteiner. 1986. Ubiquitin-lysosome conjugates. Purification and susceptibility to proteolysis. J. Biol. Chem. 261:2391-2399.
- 103 Huang, H., and A. L. Goldberg. 1997. Proteolytic activity of the ATP-dependent protease HslVU can be uncoupled from ATP hydrolysis. J. Biol. Chem. 272:21364-21372.
- 104 Hwang, B. J., K. M. Woo, A. L. Goldberg, and C. H. Chung. 1988. Protease Ti, a new ATP-dependent protease in *Escherichia coli*, contains protein-activated ATPase and proteolytic functions in distinct subunits. J. Biol. Chem. 263:8727-8734.
- 105 Johnston, J. A., E. S. Johnson, P. R. Waller, and A. Varshavsky. 1995. Methotrexate inhibits proteolysis of dihydrofolate reductase by the N-end rule pathway. J. Biol. Chem. 270:8172-8178.
- 106 Jones W.J, Paynter M.J.B., and Gupta R. 1983. Characterization of *Methanococcus maripaludis* sp. nov., a new methanogen isolated from salt-marsh sediment. Arch. Microbiol. 135:91-97.
- 107 Jones W.J, Whitman W.B, Fields R.D, and Wolfe R.S. 1983. Growth and plating efficiency of methanococci on agar media. Appl. Environ. Microbiol. 46:220-226.

- 108 Jones, W. J., J. A. Leigh, F. Mayer, C. R. Woese, and R. S. Wolfe. 1983. *Methanococcus jannaschii* sp. nov., an extremely thermophilic methanogen from a submarine hydrothermal vent. *Arch.Microbiol.* **136**:254-261.
- 109 Kanemori, M., K. Nishihara, H. Yanagi, and T. Yura. 1997. Synergistic roles of HslVU and other ATP-dependent proteases in controlling in vivo turnover of σ^{32} and abnormal proteins in *Escherichia coli*. *J.Bacteriol.* **179**:7219-7225.
- 110 Kang, M. S., B. K. Lim, I. S. Seong, J. H. Seol, N. Tanahashi, K. Tanaka, and C. H. Chung. 2001. The ATP-dependent CodWX (HslVU) protease in *Bacillus subtilis* is an N-terminal serine protease. *EMBO J.* **20**:734-742.
- 111 Karata, K., T. Inagawa, A. J. Wilkinson, T. Tatsuta, and T. Ogura. 1999. Dissecting the role of a conserved motif (the second region of homology) in the AAA family of ATPases. Site-directed mutagenesis of the ATP-dependent protease FtsH. *J.Biol.Chem.* **274**:26225-26232.
- 112 Kawarabayasi, Y., Y. Hino, H. Horikawa, S. Yamazaki, Y. Haikawa, K. Jin-no, M. Takahashi, M. Sekine, S. Baba, A. Ankai, H. Kosugi, A. Hosoyama, S. Fukui, Y. Nagai, K. Nishijima, H. Nakazawa, M. Takamiya, S. Masuda, T. Funahashi, T. Tanaka, Y. Kudoh, J. Yamazaki, N. Kushida, A. Oguchi, and H. Kikuchi. 1999. Complete genome sequence of an aerobic hyper-thermophilic crenarchaeon, *Aeropyrum pernix* K1. *DNA Res.* **6**:83-52.
- 113 Kawarabayasi, Y., M. Sawada, H. Horikawa, Y. Haidawa, Y. Hino, S. Yamamoto, M. Sekine, S. Baba, H. Kosugi, A. Hosoyama, Y. Nagai, M. Sakai, K. Ogura, R. Otsuka, H. Nakazawa, M. Takamiya, Y. Ohfuku, T. Funahashi, T. Tanaka, Y. Kudoh, J. Yamazaki, N. Kushida, A. Oguchi, K. Aoki, T. Yoshizawa, Y. Nakamura, F. T. Robb, K. Horikoshi, Y. Masuchi, H. Shizuya, and H. Kikuchi. 1998. Complete sequence and gene organization of the genome of a hyper-thermophilic archaeobacterium, *Pyrococcus horikoshii* OT3. *DNA Res.* **5**:147-155.
- 114 Kawashima, T., N. Amano, H. Koike, S. Makino, S. Higuchi, Y. Kawashima-Ohya, K. Watanabe, M. Yamazaki, K. Kanehori, T. Kawamoto, T. Nunoshiba, Y. Yamamoto, H. Aramaki, K. Makino, and M. Suzuki. 2000. Archaeal adaptation to higher temperatures revealed by genomic sequence of *Thermoplasma volcanium*. *Proc.Natl.Acad.Sci.U.S.A.* **97**:14257-14262.
- 115 Keiler, K. C., and R. T. Sauer. 1996. Sequence determinants of C-terminal substrate recognition by the Tsp protease. *J.Biol.Chem.* **271**:2589-2593.
- 116 Keiler, K. C., P. R. H. Waller, R. T. Sauer, and P. R. Waller. 1996. Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science* **271**:990-993.

- 117 Kessel, M., M. R. Maurizi, B. Kim, E. Kocsis, B. L. Trus, S. K. Singh, and A. C. Steven. 1995. Homology in structural organization between *E. coli* ClpAP protease and the eukaryotic 26S proteasome. *J.Mol.Biol.* **250**:587-594.
- 118 Kessel, M., W. Wu, S. Gottesman, E. Kocsis, A. C. Steven, and M. R. Maurizi. 1996. Six-fold rotational symmetry of ClpQ, the *E.coli* homolog of the 20S proteasome, and its ATP-dependent activator, ClpY. *FEBS Lett.* **398**:274-278.
- 119 Kihara, A., Y. Akiyama, and K. Ito. 1996. A protease complex in the *Escherichia coli* plasma membrane: HflKC (HflA) forms a complex with FtsH (HflB), regulating its proteolytic activity against SecY. *EMBO J.* **15**:6122-6131.
- 120 Kihara, A., Y. Akiyama, and K. Ito. 1998. Different pathways for protein degradation by the FtsH/HflKC membrane-embedded protease complex: an implication from the interference by a mutant form of a new substrate protein, YccA. *J.Mol.Biol.* **279**:175-188.
- 121 Kim, K. K., R. Kim, and S. H. Kim. 1998. Crystal structure of a small heat-shock protein. *Nature* **394**:595-599.
- 122 Kim, R., K. K. Kim, H. Yokota, and S. H. Kim. 1998. Small heat shock protein of *Methanococcus jannaschii*, a hyperthermophile. *Proc.Natl.Acad.Sci.U.S.A.* **95**:9129-9133.
- 123 Kim, R., S. J. Sandler, S. Goldman, H. Yokota, A. J. Clark, and S.-H. Kim. 1998. Overexpression of archaeal proteins in *Escherichia coli*. *Biotech.Lett.* **20**:207-210.
- 124 Kim, R., S. J. Sandler, S. Goldman, H. Yokota, A. J. Clark, and S. H. Kim. 1998. Overexpression of archaeal proteins in *Escherichia coli*. *Biotech.Lett.* **20**:207-210.
- 125 King, R. D., M. Saqi, R. Sayle, and M. J. Sternberg. 1997. DSC: public domain protein secondary structure predication. *Comput.Appl.Biosci.* **13**:473-474.
- 126 Kisselev, A. F., T. N. Akopian, K. M. Woo, and A. L. Goldberg. 1999. The sizes of peptides generated from protein by mammalian 26 and 20 S proteasomes. Implications for understanding the degradative mechanism and antigen presentation. *J.Biol.Chem.* **274**:3363-3371.
- 127 Kisselev, A. F., Z. Songyang, and A. L. Goldberg. 2000. Why does threonine, and not serine, function as the active site nucleophile in proteasomes? *J.Biol.Chem.* **275**:14831-14837.
- 128 Klenk, H.-P., R. A. Clayton, J. F. Tomb, O. White, K. E. Nelson, K. A. Ketchum, R. J. Dodson, M. Gwinn, E. K. Hickey, J. D. Peterson, D. L. Richardson, A. R. Kerlavage, D. E. Graham, N. C. Kyrpides, R. D. Fleischmann, J. Quackenbush, N. H. Lee, G. G. Sutton, S. Gill, E. F. Kirkness,

- B. A. Dougherty, K. McKenney, M. D. Adams, B. Loftus, and J. C. Venter. 1997. The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. *Nature* **390**:364-370.
- 129 Klotzbucher, A., E. Stewart, D. Harrison, and T. Hunt. 1996. The 'destruction box' of cyclin A allows B-type cyclins to be ubiquitinated, but not efficiently destroyed. *EMBO J.* **15**:3053-3064.
 - 130 Klumpp, M., and W. Baumeister. 1998. The thermosome: archetype of group II chaperonins. *FEBS Lett.* **430**:73-77.
 - 131 Klumpp, M., W. Baumeister, and L. O. Essen. 1997. Structure of the substrate binding domain of the thermosome, an archaeal group II chaperonin. *Cell* **91**:263-270.
 - 132 Kodama, T., K. Fukui, and K. Kometani. 1986. The initial phosphate burst in ATP hydrolysis by myosin and subfragment-1 as studied by a modified malachite green method for determination of inorganic phosphate. *J.Biochem.(Tokyo)* **99**:1465-1472.
 - 133 Koegl, M., T. Hoppe, S. Schlenker, H. D. Ulrich, T. U. Mayer, and S. Jentsch. 1999. A novel ubiquitination factor, E4, is involved in multiubiquitin chain assembly. *Cell* **96**:635-644.
 - 134 Kopp, F., K. B. Hendil, B. Dahlmann, P. Kristensen, A. Sobek, and W. Uerkevitz. 1997. Subunit arrangement in the human 20S proteasome. *Proc.Natl.Acad.Sci.U.S.A.* **94**:2939-2944.
 - 135 Kopp, F., P. Kristensen, K. B. Hendil, A. Johnsen, A. Sobek, and B. Dahlmann. 1995. The human proteasome subunit HsN3 is located in the inner rings of the complex dimer. *J.Mol.Biol.* **248**:264-272.
 - 136 Kowalski, J. M., R. M. Kelly, J. Konisky, D. S. Clark, and K. D. Wittrup. 1998. Purification and functional characterization of a chaperone from *Methanococcus jannaschii*. *Syst.Appl.Microbiol.* **21**:173-178.
 - 137 Kuo, Y. P., D. K. Thompson, A. St.Jean, R. L. Charlebois, and C. J. Daniels. 1997. Characterization of two heat shock genes from *Haloferax volcanii*: a model system for transcription regulation in the Archaea. *J.Bacteriol.* **179**:6318-6324.
 - 138 Laachouch, J. E., L. Desmet, V. Geuskens, R. Grimaud, and A. Toussiant. 1996. Bacteriophage Mu repressor as a target for the *Escherichia coli* ATP-dependent Clp protease. *EMBO J.* **15**:437-444.
 - 139 Lacalle R.A., Pulido D., Vara J., Zalacain M., and Jimenez A. 1989. Molecular analysis of the *pac* gene encoding a puromycin N-acetyl transferase from *Streptomyces alboniger*. *Gene* **79**:375-380.

- 140 **Ladapo J., and Whitman W.B.** 1990. Method for isolation of auxotrophs in the methanogenic archaeobacteria: role of the acetyl-CoA pathway of autotrophic CO₂ fixation in *Methanococcus maripaludis*. *Proc.Natl.Acad.Sci.U.S.A.* **87**:5598-5602.
- 141 **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**:680-685.
- 142 **Lanzetta, P. A., L. J. Alvarez, P. S. Reinach, and O. A. Candia.** 1979. An improved assay for nanomole amounts of inorganic phosphate. *Anal.Biochem.* **100**:95-97.
- 143 **Lenzen, C. U., D. Steinmann, S. W. Whiteheart, and W. I. Weis.** 1998. Crystal structure of the hexamerization domain of N-ethylmaleimide-sensitive fusion protein. *Cell* **94**:525-536.
- 144 **Leonhard, K., A. Stiegler, W. Neupert, and T. Langer.** 1999. Chaperone-like activity of the AAA domain of the yeast Yme1 AAA protease. *Nature* **398**:348-351.
- 145 **Levchenko, I., T. A. Baker, and L. Luo.** 1995. Disassembly of the Mu transposase tetramer by the ClpX chaperone. *Genes Dev.* **9**:2399-2408.
- 146 **Levchenko, I., M. Seidel, R. T. Sauer, and T. A. Baker.** 2000. A specificity-enhancing factor for the ClpXP degradation machine. *Science* **289**:2354-2356.
- 147 **Levchenko, I., M. Yamauchi, and T. A. Baker.** 1997. ClpX and MuB interact with overlapping regions of Mu transposase: implications for control of the transposition pathway. *Genes Dev.* **11**:1561-1572.
- 148 **Li, X., and P. Coffino.** 1993. Degradation of ornithine decarboxylase: exposure of the C-terminal target by a polyamine-inducible inhibitory protein. *Mol.Cell Biol.* **13**:2377-2383.
- 149 **Li, X., B. Stebbins, L. Hoffman, G. Pratt, M. Rechsteiner, and P. Coffino.** 1996. The N terminus of antizyme promotes degradation of heterologous proteins. *J.Biol.Chem.* **271**:4441-4446.
- 150 **Liou, A. K., and K. R. Willison.** 1997. Elucidation of the subunit orientation in CCT (chaperonin containing TCP1) from the subunit composition of CCT micro-complexes. *EMBO J.* **16**:4311-4316.
- 151 **Löwe, J., D. Stock, B. Jap, P. Zwickl, W. Baumeister, and R. Huber.** 1995. Crystal structure of the 20S proteasome from the archaeon *T. acidophilum* at 3.4 Å resolution. *Science* **268**:533-539.
- 152 **Luby-Phelps, K.** 1994. Physical properties of cytoplasm. *Curr.Opin.Cell Biol.* **6**:3-9.

- 153 **Lucero, H. A., E. W. Chojnicki, S. Mandiyan, H. Nelson, and N. Nelson.** 1995. Cloning and expression of a yeast gene encoding a protein with ATPase activity and high identity to the subunit 4 of the human 26 S protease. *J.Biol.Chem.* **270**:9178-9184.
- 154 **Macario, A. J., and E. C. de Macario.** 1999. The archaeal molecular chaperone machine. Peculiarities and paradoxes. *Genetics* **152**:1277-1283.
- 155 **Makino, Y., K. Yamano, M. Kanemaki, K. Morikawa, T. Kishimoto, N. Shimbara, K. Tanaka, and T. Tamura.** 1997. SUG1, a component of the 26 S proteasome, is an ATPase stimulated by specific RNAs. *J.Biol.Chem.* **272**:23201-23205.
- 156 **Maupin-Furlow, J. A., H. C. Aldrich, and J. G. Ferry.** 1998. Biochemical characterization of the 20S proteasome from the methanoarchaeon *Methanosarcina thermophila*. *J.Bacteriol.* **180**:1480-1487.
- 157 **Maupin-Furlow, J. A., and J. G. Ferry.** 1995. A proteasome from the methanogenic archaeon *Methanosarcina thermophila*. *J.Biol.Chem.* **270**:28617-28622.
- 158 **Maupin-Furlow, J. A., H. L. Wilson, S. J. Kaczowka, and M. S. Ou.** 2000. Proteasomes in the archaea: from structure to function. *Front.Biosci.* **5**:d837-d865.
- 159 **Maurizi, M. R., S. K. Singh, M. W. Thompson, M. Kessel, and A. Ginsburg.** 1998. Molecular properties of ClpAP protease of *Escherichia coli*: ATP-dependent association of ClpA and ClpP. *Biochemistry* **37**:7778-7786.
- 160 **Mayr, J., E. Seemüller, S. A. Muller, A. Engel, and W. Baumeister.** 1998. Late events in the assembly of 20S proteasomes. *J.Struct.Biol.* **124**:179-188.
- 161 **McGuire, M. J., and G. N. DeMartino.** 1986. Purification and characterization of a high molecular weight proteinase (macropain) from human erythrocytes. *Biochim.Biophys.Acta* **873**:279-289.
- 162 **Medigue, C., T. Rouxel, P. Vigier, A. Henaut, and A. Danchin.** 1991. Evidence for horizontal gene transfer in *Escherichia coli* speciation. *J.Mol.Biol.* **222**:851-856.
- 163 **Menon, A. S., and A. L. Goldberg .** 1987. Protein substrates activate the ATP-dependent protease La by promoting nucleotide binding and release of bound ADP. *J.Biol.Chem.* **262**:14929-14934.
- 164 **Monaco, J. J., and H. O. McDevitt.** 1984. H-2-linked low-molecular weight polypeptide antigens assemble into an unusual macromolecular complex. *Nature* **309**:797-799.

- 165 **Morgan, A., R. Dimaline, and R. D. Burgoyne.** 1994. The ATPase activity of N-ethylmaleimide-sensitive fusion protein (NSF) is regulated by soluble NSF attachment proteins. *J.Biol.Chem.* **269**:29347-29350.
- 166 **Nakai, T., T. Yasuhara, Y. Fujiki, and A. Ohashi.** 1995. Multiple genes, including a member of the AAA family, are essential for degradation of unassembled subunit 2 of cytochrome c oxidase in yeast mitochondria. *Mol.Cell Biol.* **15**:4441-4452.
- 167 **Nakayama, K. I., S. Hatakeyama, and K. Nakayama.** 2001. Regulation of the cell cycle at the G1-S transition by proteolysis of cyclin E and p27Kip1. *Biochem.Biophys.Res.Comm.* **282**:853-860.
- 168 **Nandi, D., H. Jiang, and J. J. Monaco.** 1996. Identification of MECL-1 (LMP-10) as the third IFN- γ -inducible proteasome subunit. *J.Immunol.* **156**:2361-2364.
- 169 **Neidermann, G., S. Butz, H. G. Ihlenfeldt, R. Grimm, M. Lucchiari, H. Hoschutsky, G. Jung, B. Maier, and K. Eichmann.** 1995. Contribution of proteasome-mediated proteolysis to the heirarchy of epitopes presented by major histocompatibility complex class molecules. *Immunity* **2**:289-299.
- 170 **Neuwald, A. F., L. Aravind, J. L. Spouge, and E. V. Koonin.** 1999. AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res.* **9**:27-43.
- 171 **Ng, W. V., S. P. Kennedy, G. G. Mahairas, B. Berquist, M. Pan, H. D. Shukla, S. R. Lasky, N. S. Baliga, V. Thorsson, J. Sbrogna, S. Swartzell, D. Weir, J. Hall, T. A. Dahl, R. Welti, Y. A. Goo, B. Leithauser, K. Keller, R. Cruz, M. J. Danson, D. W. Hough, D. G. Maddocks, P. E. Jablonski, M. P. Krebs, C. M. Angevine, and H. Dale.** 2000. From the cover: genome sequence of *Halobacterium* species NRC-1. *Proc.Natl.Acad.Sci.U.S.A.* **97**:12176-12181.
- 172 **Nitsch, M., M. Klumpp, A. Lupas, and W. Baumeister.** 1997. The thermosome: alternating α and β subunits within the chaperonin of the archaeon *Thermoplasma acidophilum*. *J.Mol.Biol.* **267**:142-149.
- 173 **Noll, I., S. Muller, and A. Klein.** 1999. Transcriptional regulation of genes encoding the selenium-free [NiFe]-hydrogenases in the archaeon *Methanococcus voltae* involves positive and negative control elements. *Genetics* **152**:1335-1341.
- 174 **Noll, K. M., and M. Vargas.** 1997. Recent advances in genetic analysis of hyperthermophilic archaea and bacteria. *Arch.Microbiol.* **168**:73-80.
- 175 **Orlowski, M., C. Cardozo, M. C. Hidalgo, and C. Michaud.** 1991. Regulation of the peptidylglutamyl-peptide hydrolyzing activity of the pituitary multicatalytic proteinase complex. *Biochemistry* **30**:5999-6005.

- 176 **Orlowski, M., and C. Michaud.** 1989. Pituitary multicatalytic proteinase complex. Specificity of components and aspects of proteolytic activity. *Biochemistry* **28**:9270-9278.
- 177 **Osteras, M., and U. Jenal.** 2000. Regulatory circuits in *Caulobacter*. *Curr.Opin.Microbiol.* **3**:171-176.
- 178 **Pak, M., J. R. Hoskins, S. K. Singh, M. R. Maurizi, and S. Wickner.** 1999. Concurrent chaperone and protease activities of ClpAP and the requirement for the N-terminal ClpA ATP binding site for chaperone activity. *J.Biol.Chem.* **274**:19316-19322.
- 179 **Pallen, M., B. Wren, and J. Parkhill.** 1999. 'Going wrong with confidence': misleading sequence analyses of CiaB and ClpX. *Mol.Microbiol.* **34**:195-195.
- 180 **Palombella, V. J., O. J. Rando, A. L. Goldberg, and T. Maniatis.** 1994. The ubiquitin-proteasome pathway is required for processing the NF- κ B1 precursor protein and the activation of NF- κ B. *Cell* **78**:773-785.
- 181 **Pamnani, V., T. Tamura, A. Lupas, J. Peters, Z. Cejka, W. Ashraf, and W. Baumeister.** 1997. Cloning, sequencing and expression of VAT, a CDC48/p97 ATPase homologue from the archaeon *Thermoplasma acidophilum*. *FEBS Lett.* **404**:263-268.
- 182 **Papa, F. R., A. Y. Amerik, and M. Hochstrasser.** 1999. Interaction of the Doa4 deubiquitinating enzyme with the yeast 26S proteasome. *Mol.Biol.Cell* **10**:741-756.
- 183 **Park, S. K., K. I. Kim, K. M. Woo, J. H. Seol, K. Tanaka, A. Ichihara, D. B. Ha, and C. H. Chung.** 1993. Site-directed mutagenesis of the dual translational initiation sites of the *clpB* gene of *Escherichia coli* and characterization of its gene products. *J.Biol.Chem.* **268** :20170-20174.
- 184 **Parsell, D. A., A. S. Kowal, and S. Lindquist.** 1994. *Saccharomyces cerevisiae* Hsp104 protein. Purification and characterization of ATP-induced structural changes. *J.Biol.Chem.* **269**:4480-4487.
- 185 **Parsell, D. A., A. S. Kowal, M. A. Singer, and S. Lindquist.** 1994. Protein disaggregation mediated by heat-shock protein Hsp104. *Nature* **372**:475-478.
- 186 **Patel, S., and M. Latterich.** 1998. The AAA team: related ATPases with diverse functions. *Trends Cell Biol.* **8**:65-71.
- 187 **Peloquin J.J., and Platzer E.G.** 1991. A simple, inexpensive electroelution device for the recovery of nucleic acid fragments from agarose gels. *Biotechniques* **10**:159-160.

- 188 Peters, J. M., W. W. Franke, and J. A. Kleinschmidt. 1994. Distinct 19 S and 20 S subcomplexes of the 26 S proteasome and their distribution in the nucleus and the cytoplasm. *J.Biol.Chem.* **269**:7709-7718.
- 189 Peters, J. M., J. R. Harris, A. Lustig, S. Muller, A. Engel, S. Volker, and W. W. Franke. 1992. Ubiquitous soluble Mg^{2+} -ATPase complex. A structural study. *J.Mol.Biol.* **223**:557-571.
- 190 Phipps, B. M., A. Hoffmann, K. O. Stetter, and W. Baumeister. 1991. A novel ATPase complex selectively accumulated upon heat shock is a major cellular component of thermophilic archaeobacteria. *EMBO J.* **10**:1711-1722.
- 191 Phipps, B. M., A. Hoffmann, K. O. Stetter, and W. Baumeister. 1991. A novel ATPase complex selectively accumulated upon heat shock is a major cellular component of thermophilic archaeobacteria. *EMBO J.* **10**:1711-1722.
- 192 Pickart, C. M. 1997. Targeting of substrates to the 26S proteasome. *FASEB J.* **11**:1055-1066.
- 193 Porankiewicz, J., J. Wang, and A. K. Clarke. 1999. New insights into the ATP-dependent Clp protease: *Escherichia coli* and beyond. *Mol.Microbiol.* **32**:449-458.
- 194 Ramos, P. C., J. Hockendorff, E. S. Johnson, A. Varshavsky, and R. J. Dohmen. 1998. Ump1p is required for proper maturation of the 20S proteasome and becomes its substrate upon completion of the assembly. *Cell* **92**:489-499.
- 195 Rechsteiner, M., and S. W. Rogers. 1996. PEST sequences and regulation by proteolysis. *Trends Biochem.Sci.* **21**:267-271.
- 196 Richmond, C., C. Gorbea, and M. Rechsteiner. 1997. Specific interactions between ATPase subunits of the 26 S protease. *J.Biol.Chem.* **272**:13403-13411.
- 197 Rivett, A. J. 1993. Proteasomes: multicatalytic proteinase complexes. *Biochem.J.* **291**:1-10.
- 198 Rock, K. L., C. Gramm, L. Rothstein, K. Clark, R. Stein, L. Dick, D. Hwang, and A. L. Goldberg. 1994. Inhibitors of the proteasome block the degradation of most cell proteins and the generation of peptides presented on MHC class I molecules. *Cell* **78**:761-771.
- 199 Rockel, B., J. Walz, R. Hegerl, J. Peters, D. Typke, and W. Baumeister. 1999. Structure of VAT, a CDC48/p97 ATPase homologue from the archaeon *Thermoplasma acidophilum* as studied by electron tomography. *FEBS Lett.* **451**:27-32.
- 200 Rogers, S., R. Wells, and M. Rechsteiner. 1986. Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science* **234**:364-368.

- 201 Rohrwild, M., O. Coux, H.-C. Huang, R. P. Moerschell, S. J. Yoo, J. H. Seol, C. H. Chung, A. L. Goldberg, and H. C. Huang. 1996. HslV-HslU: a novel ATP-dependent protease complex in *Escherichia coli* related to the eukaryotic proteasome. *Proc.Natl.Acad.Sci.U.S.A.* 93:5808-5813.
- 202 Rohrwild, M., G. Pfeifer, U. Santarius, S. A. Muller, H. C. Huang, A. Engel, W. Baumeister, and A. L. Goldberg. 1997. The ATP-dependent HslVU protease from *Escherichia coli* is a four-ring structure resembling the proteasome. *Nat.Struct.Biol.* 4:133-139.
- 203 Rubin, D. M., M. H. Glickman, C. N. Larsen, S. Dhruvakumar, and D. Finley. 1998. Active site mutants in the six regulatory particle ATPases reveal multiple roles for ATP in the proteasome. *EMBO J.* 17:4909-4919.
- 204 Ruepp, A., W. Graml, M. L. Santos-Martinez, K. K. Koretke, C. Volker, H. W. Mewes, D. Frishman, S. Stocker, A. N. Lupas, and W. Baumeister. 2000. The genome sequence of the thermoacidophilic scavenger *Thermoplasma acidophilum*. *Nature* 407:508-513.
- 205 Russell, S. J., S. H. Reed, W. Huang, E. C. Friedberg, and S. A. Johnston. 1999. The 19S regulatory complex of the proteasome functions independently of proteolysis in nucleotide excision repair. *Mol.Cell* 3:687-695.
- 206 Sambrook J., Fritsch E.F., and Maniatis T. 1989. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, New York.
- 207 Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc.Natl.Acad.Sci.U.S.A.* 74:5463-5467.
- 208 Saraste, M., P. R. Sibbald, and A. Wittinghofer. 1990. The P-loop--a common motif in ATP- and GTP-binding proteins. *Trends Biochem.Sci.* 15:430-434.
- 209 Schirmer, E. C., J. R. Glover, M. A. Singer, and S. Lindquist. 1996. HSP100/Clp proteins: a common mechanism explains diverse functions. *Trends Biochem.Sci.* 21:289-296.
- 210 Schumann, W. 1999. FtsH--a single-chain charonin? *FEMS Microbiol.Rev.* 23:1-11.
- 211 Seemüller, E., A. Lupas, and W. Baumeister. 1996. Autocatalytic processing of the 20S proteasome. *Nature* 382:468-470.
- 212 Seemüller, E., A. Lupas, D. Stock, J. Löwe, R. Huber, and W. Baumeister. 1995. Proteasome from *Thermoplasma acidophilum*: a threonine protease. *Science* 268:579-582.

- 213 Selkov, E., N. Maltsev, G. J. Olsen, R. Overbeek, and W. B. Whitman. 1997. A reconstruction of the metabolism of *Methanococcus jannaschii* from sequence data. *Gene* 197:GC11-26.
- 214 Sensen, C. W., R. L. Charlebois, C. Chow, I. G. Clausen, B. Curtis, W. F. Doolittle, M. Duguet, G. Erauso, T. Gaasterland, R. A. Garrett, P. Gordon, I. H. de Jong, A. C. Jeffries, C. Kozera, N. Medina, A. De Moors, O. J. van der, H. Phan, M. A. Ragan, M. E. Schenk, Q. She, R. K. Singh, and N. Tolstrup. 1998. Completing the sequence of the *Sulfolobus solfataricus* P2 genome. *Extremophiles* 2:305-312.
- 215 Sensen, C. W., H. P. Klenk, R. K. Singh, G. Allard, C. C. Chan, Q. Y. Liu, S. L. Penny, F. Young, M. E. Schenk, T. Gaasterland, W. F. Doolittle, M. A. Ragan, and R. L. Charlebois. 1996. Organizational characteristics and information content of an archaeal genome: 156 kb of sequence from *Sulfolobus solfataricus* P2. *Mol. Microbiol.* 22:175-191.
- 216 Seol, J. H., S. J. Yoo, M. S. Kang, D. B. Ha, and C. H. Chung. 1995. The 65-kDa protein derived from the internal translational start site of the *clpA* gene blocks autodegradation of ClpA by the ATP-dependent protease Ti in *Escherichia coli*. *FEBS Lett.* 377:41-43.
- 217 Seol, J. H., S. J. Yoo, K. I. Kim, M. S. Kang, D. B. Ha, and C. H. Chung. 1994. The 65-kDa protein derived from the internal translational initiation site of the *clpA* gene inhibits the ATP-dependent protease Ti in *Escherichia coli*. *J. Biol. Chem.* 269:29468-29473.
- 218 Seol, J. H., S. J. Yoo, D. H. Shin, Y. K. Shim, M. S. Kang, A. L. Goldberg, and C. H. Chung. 1997. The heat-shock protein HslVU from *Escherichia coli* is a protein-activated ATPase as well as an ATP-dependent proteinase. *Eur. J. Biochem.* 247:1143-1150.
- 219 Shanklin, J., N. D. DeWitt, and J. M. Flanagan. 1995. The stroma of higher plant plastids contain ClpP and ClpC, functional homologs of *Escherichia coli* ClpP and ClpA: an archetypal two-component ATP-dependent protease. *Plant Cell* 7:1713-1722.
- 220 She, Q., R. K. Singh, F. Confalonieri, Y. Zivanovic, G. Allard, M. J. Awayez, C. C. Chan-Weiher, I. G. Clausen, B. A. Curtis, A. De Moors, G. Erauso, C. Fletcher, P. M. Gordon, I. Heikamp-De Jong, A. C. Jeffries, C. J. Kozera, N. Medina, X. Peng, H. P. Thi-Ngoc, P. Redder, M. E. Schenk, C. Theriault, N. Tolstrup, R. L. Charlebois, W. F. Doolittle, M. Duguet, T. Gaasterland, R. A. Garrett, M. A. Ragan, C. W. Sensen, and O. J. Van Der. 2001. The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. *Proc. Natl. Acad. Sci. U.S.A.* 98:7835-7840.
- 221 Sheaff, R. J., and J. M. Roberts . 1996. End of the line: proteolytic degradation of cyclin-dependent kinase inhibitors. *Chem. Biol.* 3:869-873.

- 222 Shibatani, T., and W. F. Ward. 1995. Sodium dodecyl sulfate (SDS) activation of the 20S proteasome in rat liver. *Arch.Biochem.Biophys.* **321**:160-166.
- 223 Shieh J., Mesbah M., and Whitman W.B. 1988. Pseudoauxotrophy of *Methanococcus voltae* for acetate, leucine, and isoleucine. *J.Bacteriol.* **170**:4091-4096.
- 224 Shin, D. H., C. S. Lee, C. H. Chung, and S. W. Suh. 1996. Molecular symmetry of the ClpP component of the ATP-dependent Clp protease, an *Escherichia coli* homolog of 20 S proteasome. *J.Mol.Biol.* **262**:71-76.
- 225 Shin, D. H., S. J. Yoo, Y. K. Shim, J. H. Seol, M. S. Kang, and C. H. Chung. 1996. Mutational analysis of the ATP-binding site in HslU, the ATPase component of HslVU protease in *Escherichia coli*. *FEBS Lett.* **398**:151-154.
- 226 Shirai, Y., Y. Akiyama, and K. Ito. 1996. Suppression of *ftsH* mutant phenotypes by overproduction of molecular chaperones. *J.Bacteriol.* **178**:1141-1145.
- 227 Shotland, Y., S. Koby, D. Teff, N. Mansur, D. A. Oren, K. Tatematsu, T. Tomoyasu, M. Kessel, B. Bukau, T. Ogura, and A. B. Oppenheim. 1997. Proteolysis of the phage lambda CII regulatory protein by FtsH (HflB) of *Escherichia coli*. *Mol.Microbiol.* **24**:1303-1310.
- 228 Shrader, T. E., J. W. Tobias, and A. Varshavsky. 1993. The N-end rule in *Escherichia coli*: cloning and analysis of the leucyl, phenylalanyl-tRNA-protein transferase gene *aat*. *J.Bacteriol.* **175**:4364-4374.
- 229 Singal, D. P., M. Ye, and S. A. Quadri. 1995. Major histocompatibility-encoded human proteasome LMP2. Genomic organization and a new form of mRNA. *J.Biol.Chem.* **270** :1966-1970.
- 230 Singh, S. K., R. Grimaud, J. R. Hoskins, S. Wickner, and M. R. Maurizi. 2000. Unfolding and internalization of proteins by the ATP-dependent proteases ClpXP and ClpAP. *Proc.Natl.Acad.Sci.U.S.A.* **97**:8898-8903.
- 231 Smith, D. R., L. A. Doucette-Stamm, C. Deloughery, H. Lee, J. Dubois, T. Aldredge, R. Bashirzadeh, D. Blakely, R. Cook, K. Gilbert, D. Harrison, L. Hoang, P. Keagle, W. Lumm, B. Pothier, D. Qiu, R. Spadafora, R. Vicaire, Y. Wang, J. Wierzbowski, R. Gibson, N. Jiwani, A. Caruso, D. Bush, and J. N. Reeve. 1997. Complete genome sequence of *Methanobacterium thermoautotrophicum* Δ H: functional analysis and comparative genomics. *J.Bacteriol.* **179**:7135-7155.
- 232 Smith, P. K., R. I. Krohn, G. T. Hermanson, A. K. Mallia, F. H. Gartner, M. D. Provenzano, E. K. Fujimoto, N. M. Goeke, B. J. Olson, and D. C. Klenk. 1985. Measurement of protein using bicinchoninic acid. *Anal.Biochem.* **150**:76-85.

- 233 **Sousa, M. C., C. B. Trame, H. Tsuruta, S. M. Wilbanks, V. S. Reddy, and D. B. McKay.** 2000. Crystal and solution structures of an HslUV protease-chaperone complex. *Cell* **103**:633-643.
- 234 **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J.Mol.Biol.* **98**:503-517.
- 235 **Sowers, K. R.** 1995. Isolation of Chromosomal and Plasmid DNAs from Methanogenic Archaea, p. 369-378. *In* F. T. Robb and A. R. Place (eds.), *Archaea: A Laboratory Manual*; Methanogens. Cold Spring Harbor Press, Plainview, New York.
- 236 **Squires, C. L., S. Pedersen, B. M. Ross, and C. Squires.** 1991. ClpB is the *Escherichia coli* heat shock protein F84.1. *J.Bacteriol.* **173**:4254-4262.
- 237 **Stahlberg, H., E. Kutejova, K. Suda, B. Wolpensinger, A. Lustig, G. Schatz, A. Engel, and C. K. Suzuki.** 1999. Mitochondrial Lon of *Saccharomyces cerevisiae* is a ring-shaped protease with seven flexible subunits. *Proc.Natl.Acad.Sci.U.S.A.* **96**:6787-6790.
- 238 **Studier, F. W., A. H. Rosenberg, J. J. Dunn, and J. W. Dubendorff.** 1990. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **185**:60-89.
- 239 **Suzuki, C. K., M. Rep, J. Maarten van Dijk, K. Suda, L. A. Grivell, G. Schatz, and J. M. van Dijk.** 1997. ATP-dependent proteases that also chaperone protein biogenesis. *Trends Biochem.Sci.* **22**:118-123.
- 240 **Swaffield, J. C., J. F. Bromberg, and S. A. Johnston.** 1992. Alterations in a yeast protein resembling HIV Tat-binding protein relieve requirement for an acidic activation domain in GAL4. *Nature* **360**:768-768.
- 241 **Tagaya, M., D. W. Wilson, M. Brunner, N. Arango, and J. E. Rothman.** 1993. Domain structure of an N-ethylmaleimide-sensitive fusion protein involved in vesicular transport. *J.Biol.Chem.* **268**:2662-2666.
- 242 **Tamura, N., F. Lottspeich, W. Baumeister, and T. Tamura.** 1998. The role of tricorn protease and its amino peptidase-interacting factors in cellular protein. *Cell* **95**:637-648.
- 243 **Tamura, T., I. Nagy, A. Lupas, F. Lottspeich, Z. Cejka, G. Schoofs, K. Tanaka, R. De Mot, and W. Baumeister.** 1995. The first characterization of a eubacterial proteasome: the 20S complex of *Rhodococcus*. *Curr.Biol.* **5**:766-774.
- 244 **Tamura, T., N. Tamura, Z. Cejka, R. Hegerl, F. Lottspeich, and W. Baumeister.** 1996. Tricorn protease--the core of a modular proteolytic system. *Science* **274**:1385-1389.

- 245 **Tamura, T., N. Tamura, F. Lottspeich, and W. Baumeister.** 1996. Tricorn protease (TRI) interacting factor 1 from *Thermoplasma acidophilum* is a proline iminopeptidase. *FEBS Lett.* **398**:101-105.
- 246 **Tanaka, K.** 1998. Proteasomes: structure and biology. *J.Biochem.(Tokyo)* **123**:195-204.
- 247 **Tanaka, K., and C. Tsurumi.** 1997. The 26S proteasome: subunits and functions. *Mol.Biol.Rep.* **24**:3-11.
- 248 **Teichmann, U., L. van Dyck, B. Guiard, H. Fischer, R. Glockshuber, W. Neupert, and T. Langer.** 1996. Substitution of PIM1 protease in mitochondria by *Escherichia coli* Lon protease. *J.Biol.Chem.* **271**:10137-10142.
- 249 **Terlesky, K. C., M. J. K. Nelson, and J. G. Ferry.** 1986. Isolation of an enzyme complex with carbon monoxide dehydrogenase activity containing corrinoid and nickel from acetate-grown *Methanosarcina thermophila*. *J.Bacteriol.* **168**:1053-1058.
- 250 **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**:4673-4680.
- 251 **Thompson, M. W., S. K. Singh, and M. R. Maurizi.** 1994. Processive degradation of proteins by the ATP-dependent Clp protease from *Escherichia coli*: requirement for the multiple array of active sites in ClpB but not ATP hydrolysis. *J.Biol.Chem.* **269**:18209-18215.
- 252 **Thrower, J. A., L. Hoffman, M. Rechsteiner, and C. M. Pickart.** 2000. Recognition of the polyubiquitin proteolytic signal. *EMBO J.* **19**:94-102.
- 253 **Tobias, J. W., T. E. Shrader, G. Rocap, and A. Varshavsky.** 1991. The N-end rule in bacteria. *Science* **254**:1374-1377.
- 254 **Tomoyasu, T., J. Gamer, B. Bukau, M. Kanemori, H. Mori, A. J. Rutman, A. B. Oppenheim, T. Yura, K. Yamanaka, H. Niki, S. Hiraga, and T. Ogura.** 1995. *Escherichia coli* FtsH is a membrane-bound, ATP-dependent protease which degrades the heat-shock transcription factor σ^{32} . *EMBO J.* **14**:2551-2560.
- 255 **Trent, J. D., E. Nimmesgern, J. S. Wall, F. U. Hartl, and A. L. Horwich.** 1991. A molecular chaperone from a thermophilic archaeobacterium is related to the eukaryotic protein t-complex polypeptide-1. *Nature* **354**:490-493.
- 256 **Tu, G. F., G. E. Reid, J. G. Zhang, R. L. Moritz, and R. J. Simpson.** 1995. C-terminal extension of truncated recombinant proteins in *Escherichia coli* with a 10Sa RNA decapeptide. *J.Biol.Chem.* **270**:9322-9326.

- 257 **Tumbula D.L., T. L. Bowen, and Whitman W.B.** 1995. Methanogens: Growth of Methanogens on Solidified Medium, p. 49-56. *In* F. T. Robb and A. R. Place (eds.), *Archaea: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, New York.
- 258 **Tumbula D.L., and Whitman W.B.** 1999. Genetics of *Methanococcus*: possibilities for functional genomics in Archaea. *Mol.Microbiol.* **33**:1-7.
- 259 **Tumbula, D. L., T. L. Bowen, and W. B. Whitman.** 1997. Characterization of pURB500 from the archaeon *Methanococcus maripaludis* and construction of a shuttle vector. *J.Bacteriol.* **179**:2976-2986.
- 260 **Tumbula, D. L., R. A. Makula, and W. B. Whitman.** 1994. Transformation of *Methanococcus maripaludis* and identification of a *Pst*I-like restriction system. *FEMS Microbiol.Lett.* **121**:309-314.
- 261 **Udenfriend, S., S. Stein, P. Bohlen, W. Dairman, W. Leimgruber, and M. Weigle.** 1972. Fluorescamine: a reagent for assay of amino acids, peptides, proteins, and primary amines in the picomole range. *Science* **178**:871-872.
- 262 **Vainberg, I. E., S. A. Lewis, H. Rommelaere, C. Ampe, J. Vandekerckhove, H. L. Klein, and N. J. Cowan.** 1998. Prefoldin, a chaperone that delivers unfolded proteins to cytosolic chaperonin. *Cell* **93**:863-873.
- 263 **van Dijl, J. M., E. Kutejova, K. Suda, D. Perecko, G. Schatz, and C. K. Suzuki.** 1998. The ATPase and protease domains of yeast mitochondrial Lon: roles in proteolysis and respiration-dependent growth. *Proc.Natl.Acad.Sci.U.S.A.* **95**:10584-10589.
- 264 **van Nocker, S., Q. Deveraux, M. Rechsteiner, and R. D. Vierstra.** 1996. *Arabidopsis MBP1* gene encodes a conserved ubiquitin recognition component of the 26S proteasome. *Proc.Natl.Acad.Sci.U.S.A.* **93**:856-860.
- 265 **Varshavsky, A.** 1992. The N-end rule. *Cell* **69**:725-735.
- 266 **Varshavsky, A.** 1997. The N-end rule pathway of protein degradation. *Genes Cells* **2**:13-28.
- 267 **Verma, R., S. Chen, R. Feldman, D. Schieltz, J. Yates, J. Dohmen, and R. J. Deshaies.** 2000. Proteasomal proteomics: identification of nucleotide-sensitive proteasome-interacting proteins by mass spectrometric analysis of affinity-purified proteasomes. *Mol.Biol.Cell* **11**:3425-3439.
- 268 **Verma, R., and R. J. Deshaies.** 2000. A proteasome howdunit: the case of the missing signal. *Cell* **101**:341-344.
- 269 **Voges, D., P. Zwickl, and W. Baumeister.** 1999. The 26S proteasome: a molecular machine designed for controlled proteolysis. *Annu.Rev.Biochem.* **68**:1015-1068.

- 270 Voorhorst, W. G., R. I. Eggen, A. C. Geerling, C. Platteeuw, R. J. Siezen, and W. M. Vos. 1996. Isolation and characterization of the hyperthermostable serine protease, pyrolysin, and its gene from the hyperthermophilic archaeon *Pyrococcus furiosus*. *J.Biol.Chem.* **271**:20426-20431.
- 271 Walker, J. E., M. Saraste, M. J. Runswick, and N. J. Gay. 1982. Distantly related sequences in the α - and β -subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* **1**:945-951.
- 272 Walz, J., A. J. Koster, T. Tamura, and W. Baumeister. 1999. Capsids of tricorn protease studied by electron cryomicroscopy. *J.Struct.Biol.* **128**:65-68.
- 273 Wang, J., J. A. Hartling, and J. M. Flanagan. 1997. The structure of ClpP at 2.3 Å resolution suggests a model for ATP-dependent proteolysis. *Cell* **91**:447-456.
- 274 Wang, J., J. A. Hartling, and J. M. Flanagan. 1998. Crystal structure determination of *Escherichia coli* ClpP starting from an EM-derived mask. *J.Struct.Biol.* **124**:151-163.
- 275 Wang, W., P. M. Chevray, and D. Nathans. 1996. Mammalian Sug1 and c-Fos in the nuclear 26S proteasome. *Proc.Natl.Acad.Sci.U.S.A.* **93**:8236-8240.
- 276 Ward, J. M. 1970. The Microbial Ecology of Estuarine Methanogenesis. M.S. Thesis, University of Florida.
- 277 Wawrzynow, A., B. Banekci, and M. Zylicz. 1996. The Clp ATPases define a novel class of molecular chaperones. *Mol.Microbiol.* **21**:895-899.
- 278 Wawrzynow, A., D. Wojtkowiak, J. Marszalek, B. Banekci, M. Jonsen, B. Graves, C. Georgopoulos, and M. Zylicz. 1995. The ClpX heat-shock protein of *Escherichia coli*, the ATP-dependent substrate specificity component of the ClpP-ClpX protease, is a novel molecular chaperone. *EMBO J.* **14**:1867-1877.
- 279 Waxman, L., and A. L. Goldberg. 1982. Protease La from *Escherichia coli* hydrolyzes ATP and proteins in a linked fashion. *Proc.Natl.Acad.Sci.U.S.A.* **79**:4883-4887.
- 280 Waxman, L., and A. L. Goldberg. 1985. Protease La, the *lon* gene product, cleaves specific fluorogenic peptides in an ATP-dependent reaction. *J.Biol.Chem.* **260**:12022-12028.
- 281 Whiteheart, S. W., K. Rossnagel, S. A. Buhrow, M. Brunner, R. Jaenicke, and J. E. Rothman. 1994. N-ethylmaleimide-sensitive fusion protein: a trimeric ATPase whose hydrolysis of ATP is required for membrane fusion. *J.Cell Biol.* **126**:945-954.

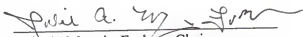
- 282 **Whitman W.B., Ankwanda E., and Wolfe R.S.** 1982. Nutrition and carbon metabolism of *Methanococcus voltae*. *J.Bacteriol.* **149**:852-863.
- 283 **Whitman W.B., Tumbula D.L., Yu J-P., and Kim W.** 1997. Development of genetic approaches for the methane-producing archaeobacterium *Methanococcus maripaludis*. *BioFactors* **6**:37-46.
- 284 **Wickner, S., S. Gottesman, D. Skowyra, J. Hoskins, K. McKenney, and M. R. Maurizi.** 1994. A molecular chaperone, ClpA, functions like DnaK and DnaJ. *Proc.Natl.Acad.Sci.U.S.A.* **91**:12218-12222.
- 285 **Wilkinson, K. D., V. L. Tashayev, L. B. O'Connor, C. N. Larsen, E. Kasperek, and C. M. Pickart.** 1995. Metabolism of the polyubiquitin degradation signal: structure, mechanism, and role of isopeptidase T. *Biochemistry* **34**:14535-14546.
- 286 **Wilson, H. L., H. C. Aldrich, and J. A. Maupin-Furlow.** 1999. Halophilic 20S proteasomes of the archaeon *Haloferax volcanii*: purification, characterization, and gene sequence analysis. *J.Bacteriol.* **181**:5814-5824.
- 287 **Wilson, H. L., M. S. Ou, H. C. Aldrich, and J. A. Maupin-Furlow.** 2000. Biochemical and physical properties of the *Methanococcus jannaschii* 20S proteasome and PAN, a homolog of the ATPase (Rpt) subunits of the eucaryal 26S proteasome. *J.Bacteriol.* **182**:1680-1692.
- 288 **Wojtkowiak, D., C. Georgopoulos, and M. Zylicz.** 1993. Isolation and characterization of ClpX, a new ATP-dependent specificity component of the Clp protease of *Escherichia coli*. *J.Biol.Chem.* **268**:22609-22617.
- 289 **Wolf, S., I. Nagy, A. Lupas, G. Pfeifer, Z. Cejka, S. A. Müller, A. Engel, R. De Mot, and W. Baumeister.** 1998. Characterization of ARC, a divergent member of the AAA ATPase family from *Rhodococcus erythropolis*. *J.Mol.Biol.* **277**:13-25.
- 290 **Woo, K. M., K. I. Kim, A. L. Goldberg, D. B. Ha, and C. H. Chung.** 1992. The heat-shock protein ClpB in *Escherichia coli* is a protein- activated ATPase. *J.Biol.Chem.* **267**:20429-20434.
- 291 **Wood A.G., Whitman W.B, and Konisky J.** 1985. A newly isolated marine methanogen harbors a small cryptic plasmid. *Arch.Microbiol.* **142**:259-261.
- 292 **Yaffe, M. B., G. W. Farr, D. Miklos, A. L. Horwich, M. L. Sternlicht, and H. Sternlicht.** 1992. TCP1 complex is a molecular chaperone in tubulin biogenesis. *Nature* **358**:245-248.
- 293 **Yan, Z., S. Fujiwara, K. Kohda, M. Takagi, and T. Imanaka.** 1997. In vitro stabilization and in vivo solubilization of foreign proteins by the beta subunit of a chaperonin from the hyperthermophilic archaeon *Pyrococcus* sp. strain KOD1. *Appl.Environ.Microbiol.* **63**:785-789.

- 294 Yoo, S. J., H. H. Kim, D. H. Shin, C. S. Lee, I. S. Seong, J. H. Seol, N. Shimbara, K. Tanaka, and C. H. Chung. 1998. Effects of the *cys* mutations on structure and function of the ATP-dependent HslVU protease in *Escherichia coli*. The Cys287 to Val mutation in HslU uncouples the ATP-dependent proteolysis by HslVU from ATP hydrolysis. *J.Biol.Chem.* **273**:22929-22935.
- 295 Yoo, S. J., J. H. Seol, D. H. Shin, M. Rohrwild, M.-S. Kang, K. Tanaka, A. L. Goldberg, C. H. Chung, and M. S. Kang. 1996. Purification and characterization of the heat shock proteins HslV and HslU that form a new ATP-dependent protease in *Escherichia coli*. *J.Biol.Chem.* **271**:14035-14040.
- 296 Yoshida, T., M. Yohda, T. Iida, T. Maruyama, H. Taguchi, K. Yazaki, T. Ohta, M. Odaka, I. Endo, and Y. Kagawa. 1997. Structural and functional characterization of homo-oligomeric complexes of alpha and beta chaperonin subunits from the hyperthermophilic archaeum *Thermococcus* strain KS-1. *J.Mol.Biol.* **273**:635-645.
- 297 Yu, R. C., P. I. Hanson, R. Jahn, and A. T. Brunger. 1998. Structure of the ATP-dependent oligomerization domain of N-ethylmaleimide sensitive factor complexed with ATP. *Nat.Struct.Biol.* **5**:803-811.
- 298 Zühl, F., E. Seemüller, R. Golbik, and W. Baumeister. 1997. Dissecting the assembly pathway of the 20S proteasome. *FEBS Lett.* **418**:189-194.
- 299 Zühl, F., T. Tamura, I. Dolenc, Z. Cejka, I. Nagy, R. De Mot, W. Baumeister, and R. De Mot. 1997. Subunit topology of the *Rhodococcus* proteasome. *FEBS Lett.* **400**:83-90.
- 300 Zwickl, P., and W. Baumeister. 1999. AAA-ATPases at the crossroads of protein life and death. *Nat.Cell Biol.* **1**:E97-8.
- 301 Zwickl, P., J. Klein, and W. Baumeister. 1994. Critical elements in proteasome assembly. *Nat.Struct.Biol.* **1**:765-770.
- 302 Zwickl, P., F. Lottspeich, and W. Baumeister. 1992. Expression of functional *Thermoplasma acidophilum* proteasomes in *Escherichia coli*. *FEBS Lett.* **312**:157-160.
- 303 Zwickl, P., D. Ng, K. M. Woo, H.-P. Klenk, and A. L. Goldberg. 1999. An archaeobacterial ATPase, homologous to ATPases in the eukaryotic 26S proteasome, activates protein breakdown by 20S proteasomes. *J.Biol.Chem.* **274**:26008-26014.

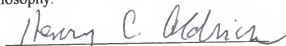
BIOGRAPHICAL SKETCH

Heather Louise Wilson was born on February 2, 1975, in Oceanside, California, to Gregory and Diane Wilson. She has two younger sisters, Holly and Heidi. In May 1996, she graduated with a Bachelor of Science degree in biology from Clearwater Christian College. In August of 1996, she entered the graduate program of the Department of Microbiology and Cell Science at the University of Florida, where she has since been working in the lab of Julie Maupin-Furlow toward completing her Ph.D degree.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.


Julie A. Maupin-Furlow, Chair
Assistant Professor of Microbiology and
Cell Science

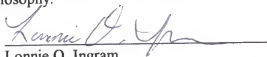
I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.


Henry C. Aldrich
Professor of Microbiology and Cell
Science

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.


Ben M. Dunn
Distinguished Professor of Biochemistry
and Molecular Biology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

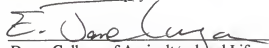

Lonnie O. Ingram
Distinguished Professor of Microbiology
and Cell Science

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.


Keelnatham T. Shanmugam
Professor of Microbiology and Cell
Science

This dissertation was submitted to the Graduate Faculty of the College of Agricultural and Life Sciences and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August 2001


Dean, College of Agricultural and Life
Sciences

Dean, Graduate School